One-point remapping of Lagrangian perturbation theory in the mildly non-linear regime of cosmic structure formation

Florent Leclercq,[1, 2, 3, a)] Jens Jasche,[1, 2] Héctor Gil-Marín,[4, 5] and Benjamin Wandelt[1, 2, 6, 7]

[1)]*Institut d'Astrophysique de Paris (IAP), UMR 7095, CNRS – UPMC Université Paris 6, Sorbonne Universités,*
98bis boulevard Arago, F-75014 Paris, France
[2)]*Institut Lagrange de Paris (ILP), Sorbonne Universités,*
98bis boulevard Arago, F-75014 Paris, France
[3)]*École polytechnique ParisTech,*
Route de Saclay, F-91128 Palaiseau, France
[4)]*Institute of Cosmology & Gravitation (ICG), University of Portsmouth,*
Dennis Sciama Building, Portsmouth PO1 3FX, UK
[5)]*Institut de Ciències del Cosmos (ICC), Universitat de Barcelona (IEEC-UB),*
Martí i Franqués 1, E-08028, Spain
[6)]*Department of Physics, University of Illinois at Urbana-Champaign,*
1110 West Green Street, Urbana, IL 61801, USA
[7)]*Department of Astronomy, University of Illinois at Urbana-Champaign,*
1002 West Green Street, Urbana, IL 61801, USA

(Dated: 5 August 2015)

On the smallest scales, three-dimensional large-scale structure surveys contain a wealth of cosmological information which cannot be trivially extracted due to the non-linear dynamical evolution of the density field. Lagrangian perturbation theory (LPT) is widely applied to the generation of mock halo catalogs and data analysis. In this work, we compare topological features of the cosmic web such as voids, sheets, filaments and clusters, in the density fields predicted by LPT and full numerical simulation of gravitational large-scale structure formation. We propose a method designed to improve the correspondence between these density fields, in the mildly non-linear regime. We develop a computationally fast and flexible tool for a variety of cosmological applications. Our method is based on a remapping of the approximately-evolved density field, using information extracted from N-body simulations. The remapping procedure consists of replacing the one-point distribution of the density contrast by one which better accounts for the full gravitational dynamics. As a result, we obtain a physically more pertinent density field on a point-by-point basis, while also improving higher-order statistics predicted by LPT. We quantify the approximation error in the power spectrum and in the bispectrum as a function of scale and redshift. Our remapping procedures improves one-, two- and three-point statistics at scales down to 8 Mpc/h.

I. INTRODUCTION

At present, observations of the three-dimensional large-scale structure (LSS) are major sources of information on the origin and evolution of the Universe. According to the current paradigm of cosmological structure formation, the presently observed structures formed via gravitational clustering of cold dark matter particles and condensation of baryonic matter in gravitational potential wells. Consequently, the large-scale matter distribution retains a memory of its formation history, enabling us to study the homogeneous as well as the inhomogeneous evolution of our Universe.

Due to non-linearities involved in the formation process, at present there exists just limited analytic understanding of structure formation in terms of perturbative expansions in Eulerian or Lagrangian representations. Both of these approaches rely on a truncated sequence of momentum moments of the Vlasov-Poisson system, completed by fluid dynamic assumptions (see e.g. Bernardeau

et al., 2002, BCGS02 hereafter, and references therein). For this reason, the validity of these approaches ceases, once the evolution of the large-scale structure enters the multi-stream regime (see e.g. Pueblas & Scoccimarro, 2009).

Nevertheless, Eulerian and Lagrangian approximations have been successfully applied to the analysis of three-dimensional density fields in regimes where they are still applicable, either at large scales or in the early Universe. Particularly, Lagrangian perturbation theory (LPT) captures significant mode-coupling information that is encoded beyond linear theory, such as large-scale flows and free-streaming, yielding three-dimensional matter distributions approximating those of full scale numerical simulations with reasonable accuracy (Moutarde *et al.*, 1991; Buchert, Melott & Weiss, 1994; Bouchet *et al.*, 1995; Scoccimarro, 1998, 2000; Scoccimarro & Sheth, 2002; Yoshisato *et al.*, 2006). Especially, second-order Lagrangian perturbation theory has been widely applied in data analysis and for fast generation of galaxy mock catalogs (e.g. PTHALOS: Scoccimarro & Sheth, 2002; Manera *et al.*, 2013; PINOCCHIO: Monaco *et al.*, 2002; Monaco, Theuns & Taffoni, 2002; Taffoni, Monaco & Theuns, 2002; Heisenberg, Schäfer & Bartelmann, 2011;

[a)]Electronic mail: florent.leclercq@polytechnique.org

Monaco *et al.*, 2013) that can be useful to estimate error bounds when analyzing observations.

Modern cosmological data analysis has an increasing demand for analytic and computationally inexpensive models providing accurate representations of the mildly non-linear regime of structure formation. Over the years, various non-linear approximations and attempts to extend the validity of LPT have been proposed. These include the spherical collapse model (Gunn & Gott, 1972), the truncated Zel'dovich approximation (Melott, Pellman & Shandarin, 1994) and models with various forms for the velocity potential (Coles, Melott & Shandarin, 1993; Munshi & Starobinsky, 1994) or the addition of a viscosity term in the Euler equation (the adhesion model, Gurbatov, Saichev & Shandarin, 1989). Analytical techniques to improve the convergence and behavior of standard perturbation theory, successfully employed in quantum field theory and statistical physics, have also been applied in the context of gravitational clustering. These include renormalized perturbation theory (Crocce & Scoccimarro, 2006), the path integral formalism (Valageas, 2007), and the renormalization group flow (Matarrese, Verde & Heavens, 1997). More recently, Tassev & Zaldarriaga (2012a,b) constructed a physical picture of the matter distribution in the mildly non-linear regime, and developed a method yielding improvements over LPT (Tassev, Zaldarriaga & Eisenstein, 2013), in particular at the scales relevant for baryon acoustic peak reconstruction (Tassev & Zaldarriaga, 2012c).

In this paper, we propose a numerically efficient method designed to improve the correspondence between approximate models and full numerical simulations of gravitational large-scale structure formation. Generally, it can be applied to any approximate model of gravitational instability, but it is especially targeted to improving Lagrangian methods. We will illustrate both these methods on fields evolved with LPT: at order one, the Zel'dovich approximation (Zel'dovich, 1970; Shandarin & Zel'dovich, 1989, ZA in the following) and second-order Lagrangian perturbation theory (2LPT in the following).

LPT and N-body density fields are visually similar, which suggests that the properties of LPT could be improved by one-to-one mapping in voxel space, following a similar line of thoughts as the "Gaussianization" idea originally proposed by Weinberg (1992, W92 hereafter) and inspired existing techniques, widely used in cosmology (Melott, 1993; Croft *et al.*, 1998; Narayanan & Weinberg, 1998; Croft *et al.*, 1999; Feng & Fang, 2000; Neyrinck, Szapudi & Szalay, 2011; Yu *et al.*, 2011, 2012; Neyrinck & Yang, 2013). The method described in this paper is based on a *remapping* of the approximately evolved particle distribution using information extracted from N-body simulations. It basically consists of replacing the one-point distribution of the approximately evolved distribution by one which better accounts for the full gravitational system. In this fashion, we adjust the one-point distribution to construct a physically more reasonable representation of the three-dimensional

matter distribution, while retaining or improving higher order statistics, described already reasonably well by the ZA (Zel'dovich, 1970; Doroshkevich, 1970; Shandarin & Zel'dovich, 1989; Buchert, 1989; Moutarde *et al.*, 1991; Yoshisato *et al.*, 2006) and by 2LPT (Moutarde *et al.*, 1991; Buchert, Melott & Weiss, 1994; Bouchet *et al.*, 1995; Scoccimarro, 1998, 2000; Scoccimarro & Sheth, 2002).

Major problems with naive approaches to remapping LPT density fields arise from minor deviations in structure types represented by LPT models and full gravity. For this reason, we compare clusters, voids, sheets, and filaments predicted by LPT and N-body simulations. Besides being of general interest for LSS data analysis, the insights gained from this comparison allow us to improve the remapping procedure.

Implementing and testing the accuracy and the regime of validity of our method is essential, and is subject of the present paper. Our study will quantify the approximation error as a function of scale in terms of a set of statistical diagnostics. From cosmographic measurements, σ_8 is known to be of order unity, which means that gravity becomes highly non-linear at some scale around 8 Mpc/h. Our method is expected to break down due to shell-crossing in LPT, at some scale larger than 8 Mpc/h. Achieving a resolution of 16 Mpc/h would already constitute substantial improvement with respect to existing methods, since non-linearities begin to affect even large-scale cosmographic measurements such as the determination of the baryon acoustic oscillations scale from galaxy surveys (about 125 Mpc/h, e.g. Eisenstein *et al.*, 2005). However, we will explore the validity of the improvement at 8 Mpc/h down to 4 Mpc/h, to see to what extent we can push the limit for possible data analysis applications into the non-linear regime. Recall that in three-dimensional LSS surveys, the number of modes usable for cosmological exploitation scales as the cube of the largest wavenumber, k^3, meaning that even minor improvements in the mildly non-linear regime would give access to much more cosmological information from existing and upcoming observations.

As will be demonstrated, this method can be used to generate realizations of density fields much faster than N-body simulations. Even though approximate, these fast realizations of mock density fields may be sufficient to model the salient features of the non-linear density field for certain applications.

This paper is structured as follows. In section II, we describe the remapping procedure for the density contrast of present-day density fields, analyze the obstacles to straightforward application and present an improved method. In section III, we apply the procedure to cosmological models using data from numerical simulations, we study the statistics of remapped fields and quantify the approximation error. We discuss our results and give our conclusions in section IV.

Corresponding LPT and N-body simulations presented in this paper were run from the same initial con-

ditions, generated at redshift $z = 63$ using second-order Lagrangian perturbation theory. The N-body simulations were run with the GADGET-2 cosmological code (Springel, Yoshida & White, 2001; Springel, 2005). Evolutions of the Zel'dovich approximation were performed with N-GENIC (Springel, 2005), and of second-order Lagrangian perturbation theory with 2LPTIC (Crocce, Pueblas & Scoccimarro, 2006). To ensure sufficient statistical significance, we used eight realizations of the same cosmology, changing the seed used to generate respective initial conditions. All computations are done after binning the density fields with a Cloud in Cell (CiC) method. The simulations contain 512^3 dark matter halos particles in a 1024 Mpc/h cubic box with periodic boundary conditions. We checked that with this setup, the power spectrum of dark matter particles agrees with the non-linear power spectrum of simulations run with higher mass resolution, provided by COSMIC EMULATOR tools (Heitmann $et\ al.$, 2009, 2010; Lawrence $et\ al.$, 2010) (deviations are at most sub-percent level for $k \lesssim 1$ (Mpc/h)$^{-1}$). Therefore, at the scales of interest of this work, $k \leq 0.4$ (Mpc/h)$^{-1}$ (corresponding to the linear and mildly non-linear regime at redshift zero), the clustering of dark matter is exactly reproduced by our set of simulations.

The cosmological parameters used are WMAP-7 fiducial values (Komatsu $et\ al.$, 2011),

$$\Omega_\Lambda = 0.728, \Omega_{\rm m} = 0.2715, \Omega_{\rm b} = 0.0455,$$
$$\sigma_8 = 0.810, h = 0.704, n_{\rm s} = 0.967. \tag{1}$$

II. REMAPPING LAGRANGIAN PERTURBATION THEORY

In this section, we discuss the remapping procedure and apply it to cosmological density fields evolved with LPT. Naively following the approach of W92 for present-day density fields yields the procedure described in section II A. This approach is not entirely satisfactory and we analyze the reasons for its shortcomings in section II B. In consequence, we propose a improvement of the remapping procedure in section II C. The properties of the remapping function are examined in section II D.

A. Remapping procedure

In this section, we describe the remapping algorithm used to go from a low-redshift realization of a density field evolved with LPT to one evolved with full N-body gravitational dynamics. Note that both fields obey the same initial conditions but are evolved by different physical models.

Density fields are defined on Cartesian grids of cubic voxels. Linear gravitational evolution exactly maintains the relative amplitude of fluctuations in different voxels. Due to mode coupling, positive and negative fluctuations grow at different rates in the non-linear regime, but even non-linear evolution tends to preserve the $rank\ order$ of the voxels, sorted by density.

The one-point probability distribution functions (PDF) and the cumulative distribution functions (CDF) of the final density fields, evolved with either LPT or full N-body gravitational dynamics, exhibit similar, but not identical shapes. This result suggests a way to improve the approximation with information extracted from the N-body simulation: maintain the rank order of the voxels, but reassign densities so that the two CDFs match. The method therefore resembles the "Gaussianization" procedure proposed by W92, an attempt to reconstruct the initial conditions of a density field from its final CDF.

Let $\mathcal{P}_{\rm LPT}$ and $\mathcal{P}_{\rm Nbody}$ denote the probability distribution functions for the density contrast in the LPT and in the full N-body density fields, respectively. Let $\mathcal{C}_{\rm LPT}$ and $\mathcal{C}_{\rm Nbody}$ be their integrals, the cumulative distribution functions. $\mathcal{C}_{\rm LPT}(\delta_{\rm LPT})$ is the $fractional\ rank$ for $\delta_{\rm LPT}$ i.e. the probability that the density contrast at a given voxel is smaller than $\delta_{\rm LPT}$, $\mathcal{P}_{\rm LPT}(\delta \leq \delta_{\rm LPT})$, and the analogous for the N-body field. The remapping procedure works as follows. A voxel with rank order $\delta_{\rm LPT}$ is assigned a new density $\delta_{\rm Nbody}$ such that

$$\mathcal{C}_{\rm LPT}(\delta_{\rm LPT}) = \mathcal{C}_{\rm Nbody}(\delta_{\rm Nbody}) \tag{2}$$

(also see figure 1 for a schematic outline of this method). The left panels of figure 1 show $\mathcal{P}_{\rm LPT}$ (dashed curves) and the corresponding cumulative distributions, $\mathcal{C}_{\rm LPT}$ (solid curves). On the right panels, the dotted curves represent the PDF of the corresponding N-body realization, $\mathcal{P}_{\rm Nbody}$. Remapping assigns to a voxel with density contrast $\delta_{\rm LPT}$ and fractional rank $\varepsilon = \mathcal{C}_{\rm LPT}(\delta_{\rm LPT})$ the value of $\delta_{\rm Nbody}$ that would have the same fractional rank in the N-body distribution.

Since $\mathcal{C}_{\rm Nbody}$ contains exactly the same information as $\mathcal{P}_{\rm Nbody}$, the remapping procedure imposes the one-point distribution taken from the N-body-evolved density field while maintaining the rank order of the LPT-evolved density field. In other words, only the weight of underdensities and overdensities is modified, while their locations remain unchanged (see also figure 2). In this fashion, we seek to adjust the density field while maintaining higher-order statistics provided by LPT with reasonable accuracy. We checked numerically that mass is always conserved in this procedure.

B. Comparison of structure types in LPT and in N-body dynamics

We implemented the remapping procedure described in the previous section and checked that we are able to modify LPT density fields so as to correctly reproduce the one-point distribution of a full N-body simulation. However, we experience a large-scale bias in the power spectrum, namely the amplitude of the spectrum of the remapped fields is slightly too high. Generally,
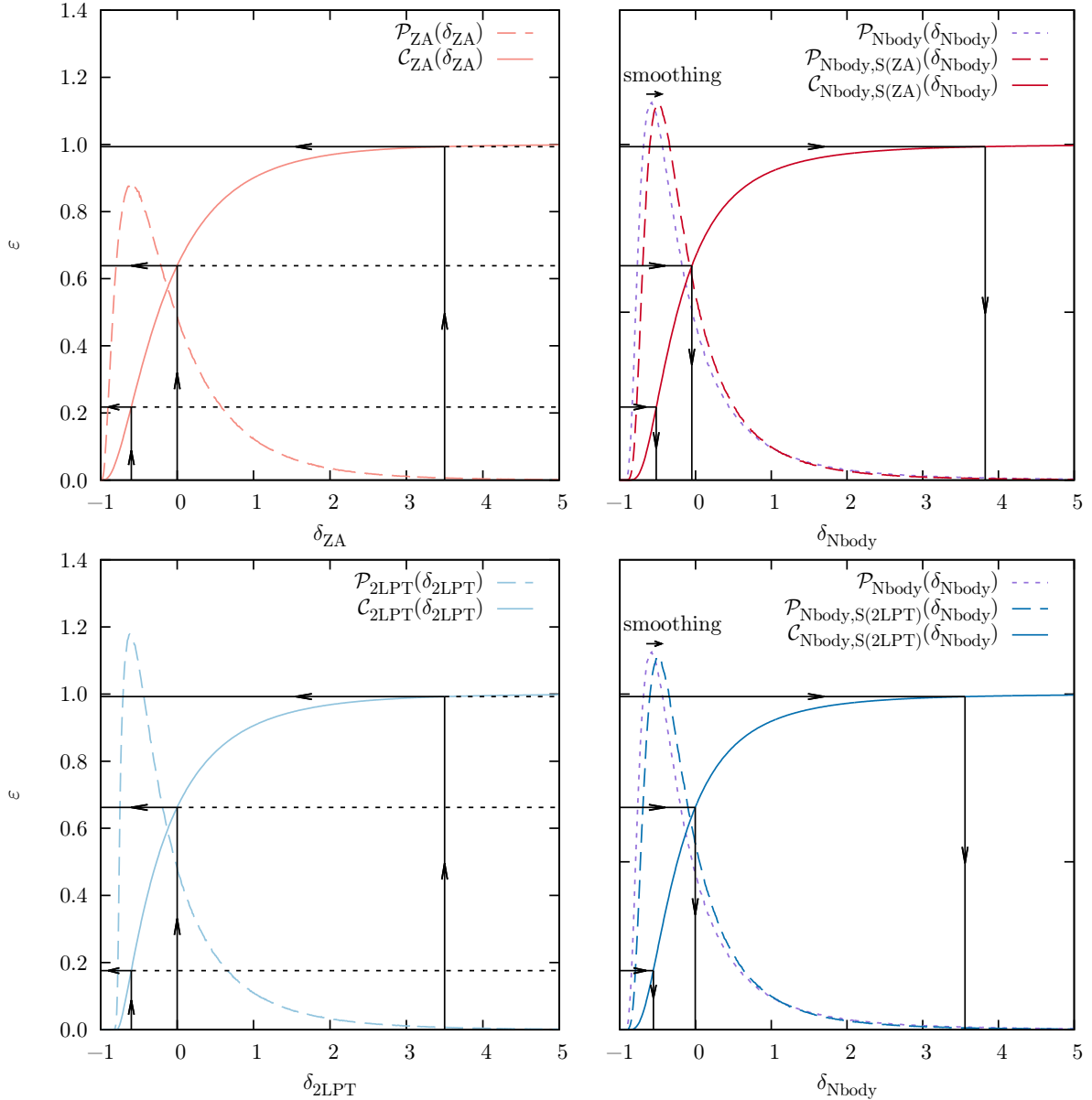
FIG. 1. A graphical illustration of the improved remapping procedure at redshift zero, for the ZA (upper panel) and 2LPT (lower panel). On the right panels, the dotted purple curves are the probability distribution function for the density contrast in full N-body dynamics. The first step of the procedure is to smooth the N-body density field using the transfer function given by eq. (4), which yields slightly different PDFs (dashed dark red and dark blue curves on the right panels). On the left panels, the dashed curves are the PDFs for the density contrast in the final density field, evolved from the same initial conditions with LPT (ZA: red line and 2LPT: blue line). The PDFs are computed on a 8 Mpc/h mesh (128^3-voxel grid) using the full statistics from eight realizations of 512^3-particles in a 1024 Mpc/h box with periodic boundary conditions. The solid curves are their respective integrals, the cumulative distribution functions. The second step is remapping, which assigns a voxel with density contrast δ_{LPT} and fractional rank ε the value of δ_{Nbody} that would have the same fractional rank in the smoothed N-body distribution (eq. (5)). This remapping is illustrated for 2LPT with three sample points: $\delta_{\mathrm{2LPT}} = -0.60$ maps to $\delta_{\mathrm{Nbody}} = -0.56$, $\delta_{\mathrm{2LPT}} = 0.00$ maps to $\delta_{\mathrm{Nbody}} = -0.01$, and $\delta_{\mathrm{2LPT}} = 3.50$ maps to $\delta_{\mathrm{Nbody}} = 3.56$. The remapping procedure imposes the one-point distribution of the smoothed N-body field while maintaining the rank order of the LPT-evolved density fields. The last step is an increase of small-scale power in the remapped distribution using the reciprocal transfer function, eq. (6).
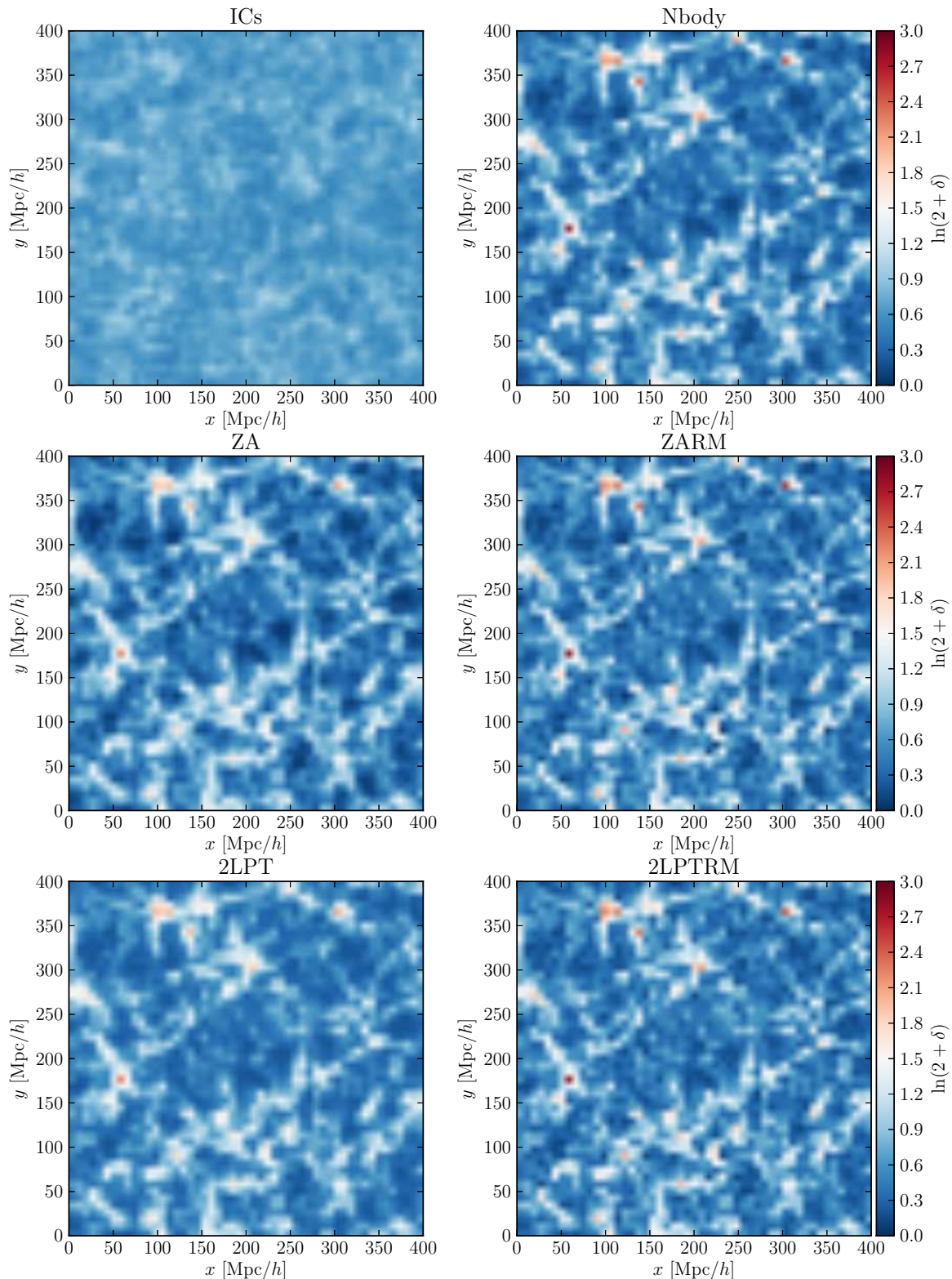
FIG. 2. Redshift-zero density contrast on a 128^2-pixel slice of a 512^3-particles realization in a 1024 Mpc/h box with periodic boundary conditions. For clarity, the slice is limited to a square of 400 Mpc/h side, and the quantity shown is the log-density field, $\ln(2 + \delta)$. For comparison with the initial conditions, the density field at high redshift ($z = 5$) is shown in the top left corner. The redshift-zero density fields are determined using, from top to bottom and from left to right: a full N-body simulation, the Zel'dovich approximation, alone (ZA) and after remapping (ZARM), second-order Lagrangian perturbation theory, alone (2LPT) and after remapping (2LPTRM). The remapping and the transfer function operations are performed on a 128^3-voxel grid, corresponding to a mesh size of 8 Mpc/h.

a non-linear transformation in voxel space can change the variance of a field. This is consistent with the findings of Weinberg (W92), who found a similar effect in his reconstructions of the initial conditions, and who addressed the issue by rescaling the density field in Fourier space. However, such an approach will generally destroy the remapped one-point distribution, and may even further lead to Gibbs ringing effects which will make the remapped field unphysical.

The bias indicates a stronger clustering in the remapped density field. Since remapping is a local operation in voxel space, this large-scale bias means that erroneous remapping of small-scale structures affects the clustering at larger scales. To identify the cause of this bias, we performed a study of differences in structure types in density fields predicted by LPT and N-body simulations. We employed the web-type classification algorithm proposed by Hahn *et al.* (2007), which relies on estimating the eigenvalues of the Hessian of the gravitational potential. In particular, this algorithm dissects the voxels into four different web types (voids, sheets, filaments and clusters). With this analysis, we wanted to understand the effect of remapping in different density and dynamical regimes of the LSS. In particular, overdense clusters are objects in the strongly non-linear regime, far beyond shell-crossing, where predictions of LPT fail, while underdense voids are believed to be better apprehended (e.g. BCGS02).

As an indicator of the mismatch between the volume occupied by different structure types in LPT and N-body dynamics, we used the quantities γ_t defined by

$$\gamma_t \equiv \frac{N_t^{\mathrm{LPT}} - N_t^{\mathrm{Nbody}}}{N_t^{\mathrm{Nbody}}}, \qquad (3)$$

where t is one of the four structure types (voids, sheets, filaments, clusters), and N_t^{LPT} and N_t^{Nbody} are the numbers of voxels flagged as belonging to a structure of type t, in corresponding LPT and in N-body realizations, respectively.

In figure 3, we plot γ_t as a function of the voxel size used to define the density fields. γ_t is positive for clusters and voids, and negative for sheets and filaments, meaning that too large cluster and void regions are predicted in LPT, at the detriment of sheets and filaments. More specifically, LPT predicts fuzzier halos than N-body dynamics, and incorrectly predicts the surroundings of voids as part of them. This result indicates that even though LPT and N-body fields look visually similar, there are crucial differences in the representation of structure types. As demonstrated by figure 3, this mismatch increases with increasing resolution. This effect is of general interest when employing LPT in LSS data analysis.

In figure 4, we plot the large-scale bias observed in remapped fields obtained with the procedure of section II A as a function of γ_t, for various resolutions. A strong correlation is observed between the bias and the mismatch in the volume occupied by different structure types. The difference in the prediction of the volume of extended objects is the cause of the bias: in clusters and in voids, remapping enhances a larger volume than should be enhanced, which yields on average a stronger clustering in the box.

C. Improvement of the remapping procedure

In the previous section, we noted that because too much volume of LPT density fields is mapped to the tails of the N-body one-point distribution (dense clusters and deep voids), the average two-point correlations are larger after remapping. Given this result, minor modifications of the procedure described in section II A can solve the problem. We now propose to improve the remapping procedure of W92 for present-day density fields, in a similar fashion as in Narayanan & Weinberg (1998), combining it with transfer function techniques to deal with the mildly non-linear modes as in Tassev & Zaldarriaga (2012a,b). Our method works as follows.

1. We degrade the N-body density field to the same degree of smoothness as the LPT density field, by multiplying the Fourier modes of the density field by the transfer function

$$T(k) \equiv \sqrt{\frac{P_{\mathrm{LPT}}(k)}{P_{\mathrm{Nbody}}(k)}}. \qquad (4)$$

   This steps yields a new density field, noted Nbody,S(LPT), whose power spectrum matches that of the LPT density field.

2. We remap the LPT density field in the fashion described in section II A, but using as a reference the CDF of the smoothed density field, $\mathcal{C}_{\mathrm{Nbody,S(LPT)}}$, instead of the full N-body density field (see figure 1). The remapping condition, eq. (2), now reads

$$\mathcal{C}_{\mathrm{LPT}}(\delta_{\mathrm{LPT}}) = \mathcal{C}_{\mathrm{Nbody,S(LPT)}}(\delta_{\mathrm{Nbody}}). \qquad (5)$$

3. We increase the power of small scales modes in the remapped distribution to the same value as in a full N-body simulation, using the reciprocal of the transfer function (4), namely

$$T^{-1}(k) = \sqrt{\frac{P_{\mathrm{Nbody}}(k)}{P_{\mathrm{LPT}}(k)}}. \qquad (6)$$

This procedure cures the large-scale bias issue experienced with the simple implementation of the remapping described in section II A, without requiring any prior knowledge on the corresponding N-body simulation. As we will demonstrate in section III, it yields improvement of one-, two- and three-point statistics of LPT.
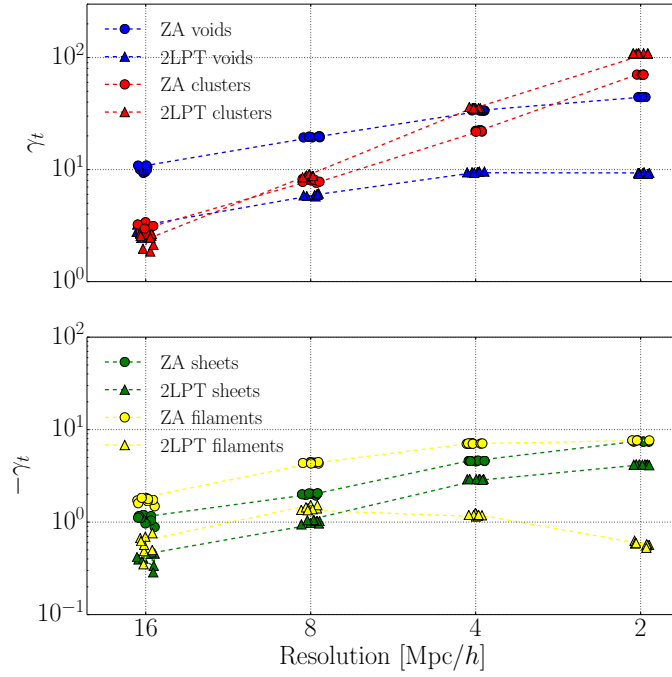
FIG. 3. Relative volume fraction of voids, sheets, filaments and clusters predicted by LPT, compared to N-body simulations, as a function of the resolution used for the definition of the density fields. The points are sightly randomized on the x-axis for clarity. The estimators γ_t are defined by eq. (3). Eight realizations of the ZA (circles) and 2LPT (triangles) are compared to the corresponding N-body realization, for various resolutions. The volume fraction of incorrectly predicted structures in LPT generally increases with increasing resolution.
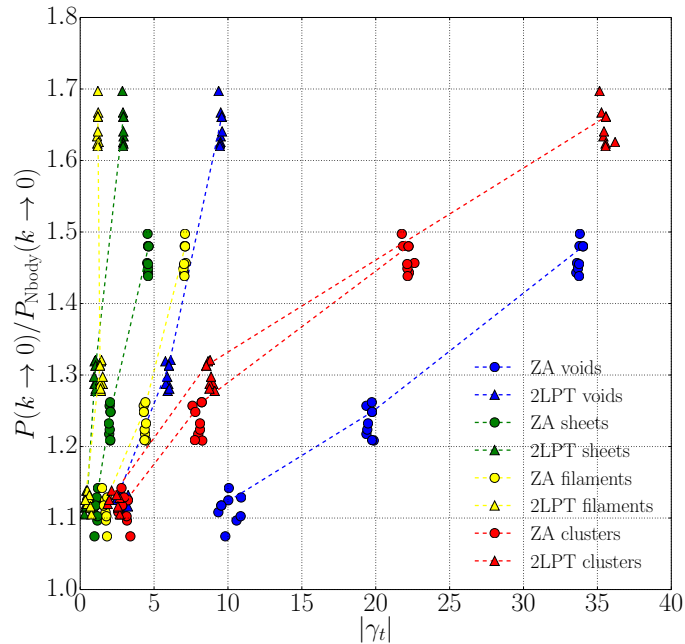
FIG. 4. The large-scale bias in the power spectrum of density fields, remapped with the procedure described in section II A, as a function of the mismatch between structure types. The estimators γ_t are defined by eq. (3). Eight realizations of the ZA (circles) and 2LPT (triangles) are compared to the corresponding N-body realization, for various resolutions (from bottom to top: 16 Mpc/h, 8 Mpc/h, 4 Mpc/h). The large-scale bias in the power spectra of remapped fields is strongly correlated to the volume fraction of structures incorrectly predicted by LPT.

D. Remapping function and transfer function

Since $\mathcal{C}_{\mathrm{LPT}}$ and $\mathcal{C}_{\mathrm{Nbody,S(LPT)}}$ are monotonically increasing functions, there is no ambiguity in the choice of δ_{Nbody}, and this procedure defines a *remapping function* f such that

$$\delta_{\mathrm{LPT}} \mapsto \delta_{\mathrm{Nbody}} = \mathcal{C}_{\mathrm{Nbody,S(LPT)}}^{-1}(\mathcal{C}_{\mathrm{LPT}}(\delta_{\mathrm{LPT}})) \equiv f(\delta_{\mathrm{LPT}}). \tag{7}$$

Establishing a remapping function f requires knowledge of both LPT and N-body density field statistics. Ideally, several realizations with different initial conditions should be combined in order to compute a precise remapping function. Indeed, a limited amount of available N-body simulations results in a lack of statistics and hence uncertainties for the remapping procedure in the high density regime. However, this effect is irrelevant from a practical point of view, since these high density events are very unlikely and affect only a negligible number of voxels. As a consequence this uncertainty will only affect to sub-percent level the usual statistical summaries of the density field. Note that in any case, if desired, the accuracy of the remapping function in the high density regime can be trivially enhanced by enlarging the size or number of N-body simulations used for its construction. For the analysis presented in this paper, the remapping functions have been computed using the full statistics from eight realizations of 512^3 particles in a 1024 Mpc/h box.

Note that once the relationship between the statistical behavior of the LPT fields and the full non-linear field is known, this procedure can be used on LPT realizations without the need of evolving corresponding N-body simulations. More specifically, the remapping function f (eq. (7)) and the transfer function T (eq. (4)) can be tabulated and stored, then used for the fast construction of a large number of large-scale structure density fields. Since producing LPT realizations is computationally faster than running full gravitational simulations by a factor of several hundreds, our method can be used to produce a large set of N-body-like realizations in a short time.

Some remapping functions are presented in figure 5. In each panel, the solid curves represent the remapping function f_z at redshift z, computed with the LPT and N-body simulations. The dashed black line shows the identity function. We checked that the remapping function converges to the identity function with increasing redshift, as expected. Critical values where the remapping function crosses the identity function are identified. Between these critical values, remapping either increases or decreases the local density.

The PDFs for the density contrast are evaluated on a grid after a CiC assignment of particles. This means that the remapping function a priori depends on the size of voxels. The problem of choosing a voxel size for the computation of the remapping function is linked to the more general problem of choosing a mesh for the CiC

evaluation of density. Choosing too coarse a binning will result in an underestimation of the clustering of particles, whereas choosing too fine a binning will also result in artifacts in overdensities (some voxels may be empty due to their too small size). The right choice of voxel size for the evaluation of the remapping function is the one giving the best evaluation of the density contrast. This choice has to be made depending on the desired application of the remapped data.

The remapping function describes how the PDF for the density contrast is affected by non-linear structure formation. For this reason, it depends on the nature of the gravitational interaction, as described by LPT and by full N-body dynamics, but weakly depends on the detail of the cosmological parameters. We checked the cosmology-dependence of the remapping function in simulations with the dark matter and baryon density in the Universe, Ω_{m} and Ω_{b}, varying the fiducial values (eq. (1)) by $\pm 3\sigma$ (still assuming a flat Universe):

$$\Omega_\Lambda = 0.750, \Omega_{\mathrm{m}} = 0.2494, \Omega_{\mathrm{b}} = 0.0428,$$
$$\sigma_8 = 0.810, h = 0.704, n_{\mathrm{s}} = 0.967; \tag{8}$$

$$\Omega_\Lambda = 0.700, \Omega_{\mathrm{m}} = 0.2992, \Omega_{\mathrm{b}} = 0.0488,$$
$$\sigma_8 = 0.810, h = 0.704, n_{\mathrm{s}} = 0.967. \tag{9}$$

Even for these models notably far from the fiducial values, we found that the remapping function almost perfectly overlaps that of our main analysis, for the density range $\delta \in [-1; 5]$, containing typically 98 to 99% of the voxels. We found a difference of less than 5% for $\delta = 5$ (see the left panel of figure 6).

The transfer function used in steps 1 and 3 of the improved procedure also exhibits very weak redshift-dependence, with deviations limited to a few percents at the smallest scales of interest of this work ($k \approx 0.4$ (Mpc/h)$^{-1}$, see the right panel of figure 6).

III. STATISTICS OF REMAPPED FIELDS

In this section, we discuss the validity of the improved remapping procedure described in section II C, by studying the correlators of the remapped field in comparison to the input LPT and N-body fields. The remapping procedure based on the Eulerian density contrast essentially replaces the LPT one-point function by that of the smoothed N-body-evolved field. Since the position and shape of structures is left unchanged, we expect the higher-order correlators of the density field to be respected by the remapping procedure. Of particular interest is to check how remapping affects higher-order statistics and if possible improvements could be exploited in data analysis or artificial galaxy survey applications.

We implemented a numerical algorithm that computes and analyzes a remapped density field. The procedure can be divided in three steps:
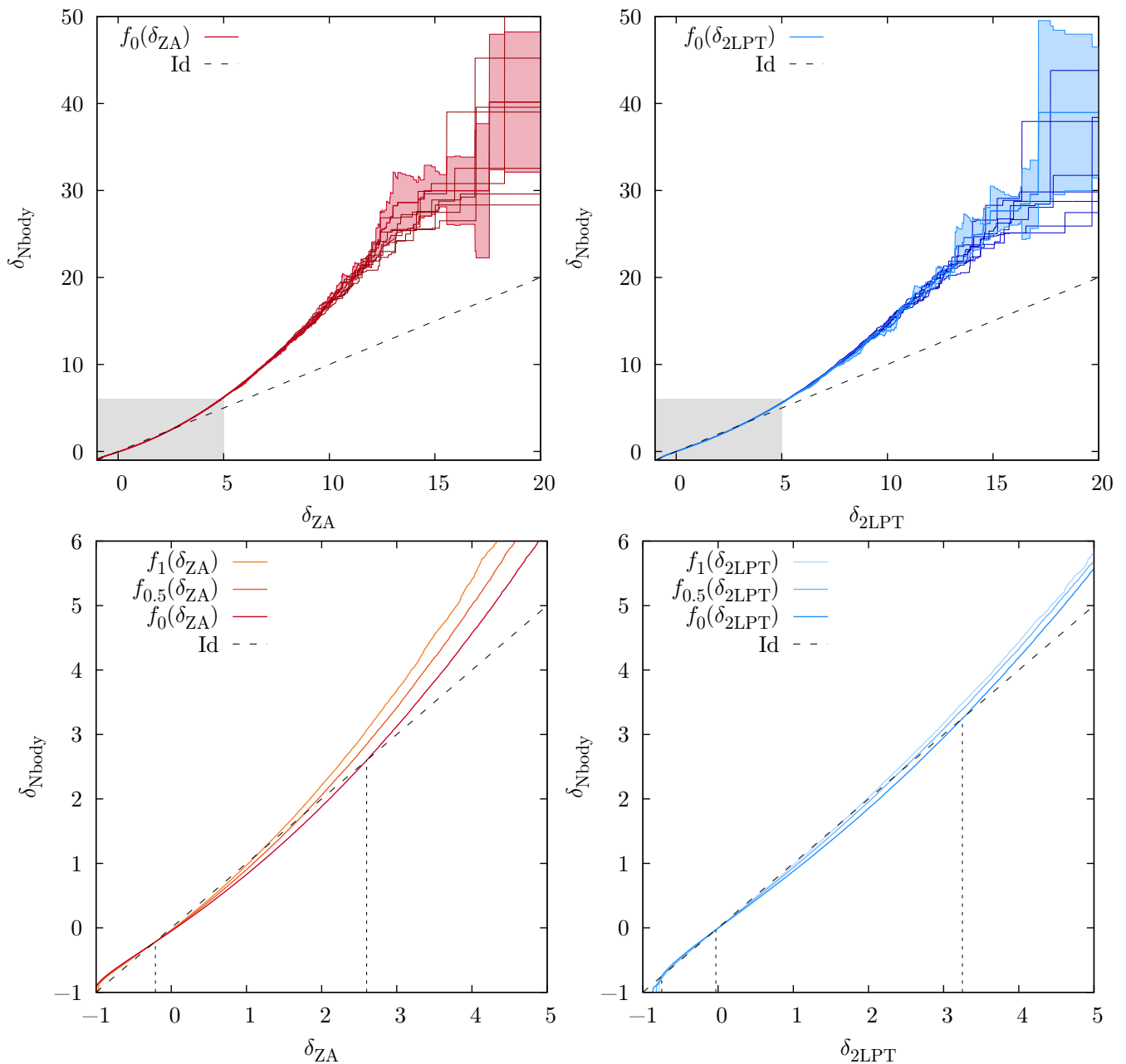
FIG. 5. The remapping functions from LPT to smoothed N-body density fields, for the ZA (left panel) and 2LPT (right panel), all computed on a 8 Mpc/h mesh. The precise redshift-zero remapping functions f_0 (red and blue solid curves) have been computed using the full statistics from eight realizations (darker red and blue solid curves). The error bars shown are the 1-σ dispersion among the eight runs with reference to the full remapping function. The lower plots show the detail of the shaded area, in a density range containing most of the voxels. The redshift-dependence of the remapping function f_z is shown for $z = 1$, $z = 0.5$ and $z = 0$. The dashed line shows the identity function. Critical values of δ_{LPT} for which remapping does not change the local density are identified.

1. We take as input two cosmological density fields, evolved from the same initial conditions with LPT (ZA or 2LPT) and with full N-body dynamics, and estimate the one-point statistics (PDF and CDF for δ) and the transfer function for this particular realization. We repeat this step for the eight real-

izations used in our analysis.

2. We take as input all the one-point statistics computed with individual realizations, and we compute a precise remapping function using the full statistics of all available realizations, as described in section II A. The transfer function used as a refer-
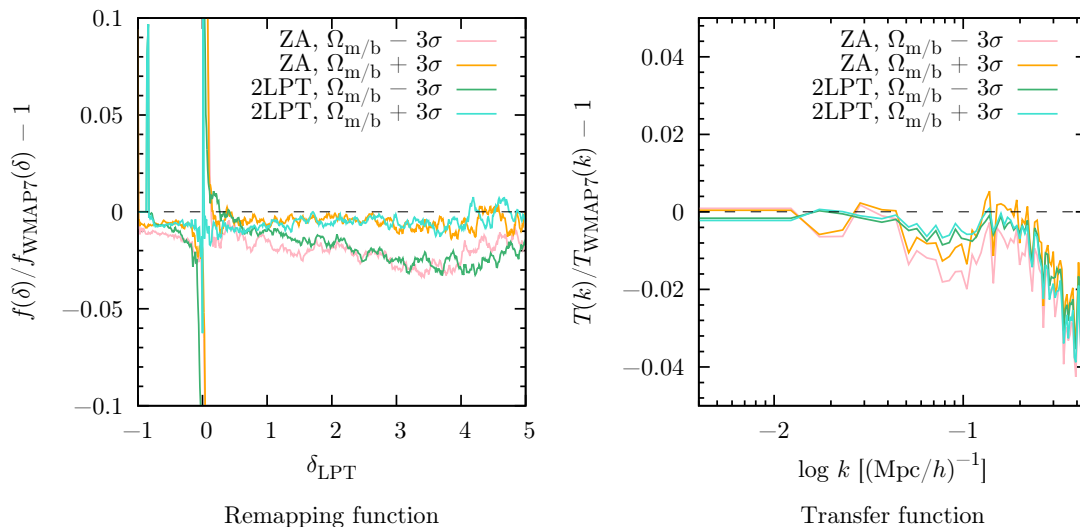
FIG. 6. Relative deviations of the remapping function (left panel) and of the transfer function (right panel) for varying cosmological parameters (eq. (8) and eq. (9)), with respect to their behaviour in a fiducial cosmology (eq. (1)).

ence is the mean of all available realizations. At this point, the remapping function and the transfer function can be tabulated and stored for later use, and N-body simulations are no longer required.

3. For each realization, we remap the density field using the improved three-step procedure described in section II C and we analyze its correlation functions.

Our algorithm provides the one-point (section III A) and two-point (section III B) statistics. We used the code described in Gil-Marín *et al.* (2011, 2012) to study the three-point statistics (section III C). The results are presented below.

A. One-point statistics

The remapping procedure, described in section II, is essentially a replacement of the cumulative distribution function of the density contrast δ of the input LPT-evolved field, $\mathcal{C}_{\mathrm{LPT}}$, by that of the reference N-body-evolved field after smoothing, $\mathcal{C}_{\mathrm{Nbody,S(LPT)}}$. After having applied the remapping procedure, we recomputed the PDF of the remapped field and verified that it matches that of the fiducial field as a sanity check.

Remapping and rescaling the density modes alters local density values but positions of structures remain unchanged. It is therefore important to check that remapping visually alters the LPT-evolved distribution in such a way that structures resemble more their N-body evolved counterparts. Figure 2 shows a slice of the density contrast δ, measured at redshift zero, on a 128^2-pixel sheet of a 512^3-particles realization in a 1024 Mpc/h box. The corresponding mesh size is 8 Mpc/h. Visually,

remapped fields (ZARM and 2LPTRM) are closer to the full N-body result than their originals (ZA and 2LPT), with plausible particle distribution.

Since the improved remapping procedure involves a rescaling of the density modes in Fourier space (step 3), the PDF for the density contrast of the remapped fields is not guaranteed to be correct by construction, as would be the case with a naive remapping (section II A). Therefore, the one-point distribution has to be carefully checked at this point. In figure 7, we plot the probability distribution function for the density contrast at redshift zero for N-body simulations and the approximately evolved fields: with the ZA and 2LPT alone, and after remapping (ZARM and 2LPTRM). It can be observed that the peaks of the PDFs get closer to the reference set by N-body dynamics and that the PDF of remapped fields accurately follows that of full gravitational dynamics for $\delta > 0$. The procedure being successful on average for one-point statistics and accurate for the most common events in overdensities, we expect the number count of objects such as clusters predicted by LPT to be made more robust by our procedure.

B. Two-point statistics

1. Power spectrum

We measured the power spectrum of dark matter particles, displaced according to each approximation and assigned to cells with a CiC scheme, for different mesh sizes. Power spectra were measured from theses meshes, with a correction for aliasing effects (Jing, 2005). Redshift-zero results computed on a 8 Mpc/h mesh are presented in figure 8. There, the dashed line corresponds to the

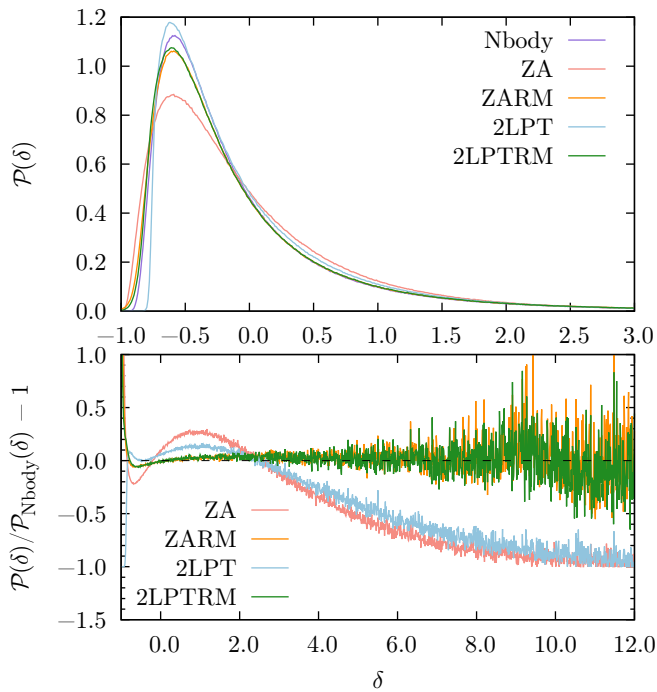FIG. 7. *Upper panel.* Redshift-zero probability distribution function for the density contrast δ, computed from eight 1024 Mpc/h-box simulations of 512^3 particles. The particle distribution is determined using: a full N-body simulation (purple curve), the Zel'dovich approximation, alone (ZA, light red curve) and after remapping (ZARM, orange curve), second-order Lagrangian perturbation theory, alone (2LPT, light blue curve) and after remapping (2LPTRM, green curve). *Lower panel.* Relative deviations of the same PDFs with reference to N-body simulation results. Note that, contrary to standard LPT approaches, remapped fields follow the one-point distribution of full N-body dynamics in an unbiased way, especially in the high density regime.

FIG. 8. Redshift-zero dark matter power spectra in a 1024 Mpc/h simulation, with density fields computed with a mesh size of 8 Mpc/h. The particle distribution is determined using: a full N-body simulation (purple curve), the Zel'dovich approximation, alone (ZA, light red curve) and after remapping (ZARM, orange curve), second-order Lagrangian perturbation theory, alone (2LPT, light blue curve) and after remapping (2LPTRM, green curve). The dashed black curve represents $P_{\mathrm{NL}}(k)$, the theoretical power spectrum expected at $z = 0$. Both ZARM and 2LPTRM show increased power in the mildly non-linear regime compared to ZA and 2LPT (at scales corresponding to $k \gtrsim 0.1$ (Mpc/h)$^{-1}$ for this redshift), indicating an improvement of two-point statistics with the remapping procedure.

theoretical, non-linear power spectrum expected, computed with Cosmic Emulator tools (Heitmann *et al.*, 2009, 2010; Lawrence *et al.*, 2010). A deviation of full N-body simulations from this theoretical prediction can be observed at small scales. This discrepancy is a gridding artifact, completely due to the finite mesh size used for the analysis. The relative deviations of various power spectra with reference to the density field computed with a full N-body simulation are presented in figures 9 and 10. In all the plots, the error bars represent the dispersion of the mean among eight independent realizations.

At high redshift ($z > 1$), we found no notable difference between the power spectrum of matter evolved with full N-body dynamics and that of LPT-remapped distributions. This indicates that our remapping procedure is increasingly successful as we go backwards in time towards the initial conditions, where LPT gives accurate results.

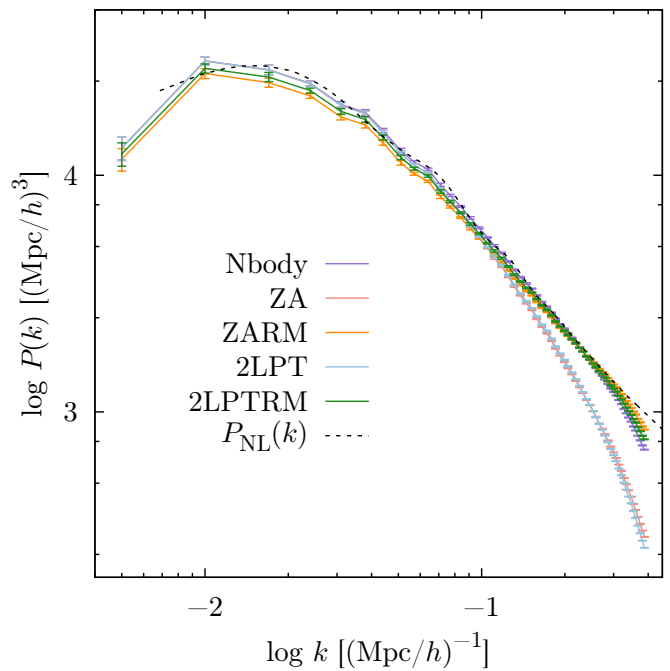At low redshift, the power spectrum shape of remapped LPT fields is closer to the shape of the full non-linear

power spectrum, turning down at smaller scales than the LPT power spectra. In particular, LPT fields exhibit more small-scale correlations after remapping, illustrating the success of our procedure in the mildly non-linear regime of large-scale structure formation.

Contrary to the density fields obtained via a naive remapping approach, whose power spectra exhibit a positive artificial offset at large scales as discussed in section II B, the fields obtained with the improved procedure have correct two-point statistics at all scales for coarse grids (down to 8 Mpc/h). For finer grids, a negative large-scale bias appears in the power spectrum, meaning that we have suppressed too much small-scale power in the N-body field in the first step of our procedure, which propagates to large scales with remapping. Comparing the panels of figures 9 and 10, it can be observed that this effect is suppressed at higher redshift and for coarser binning. We found that a voxel size of 8 Mpc/h is the best compromise, with a large-scale power spectrum compatible with the error bars and clear improvement at small
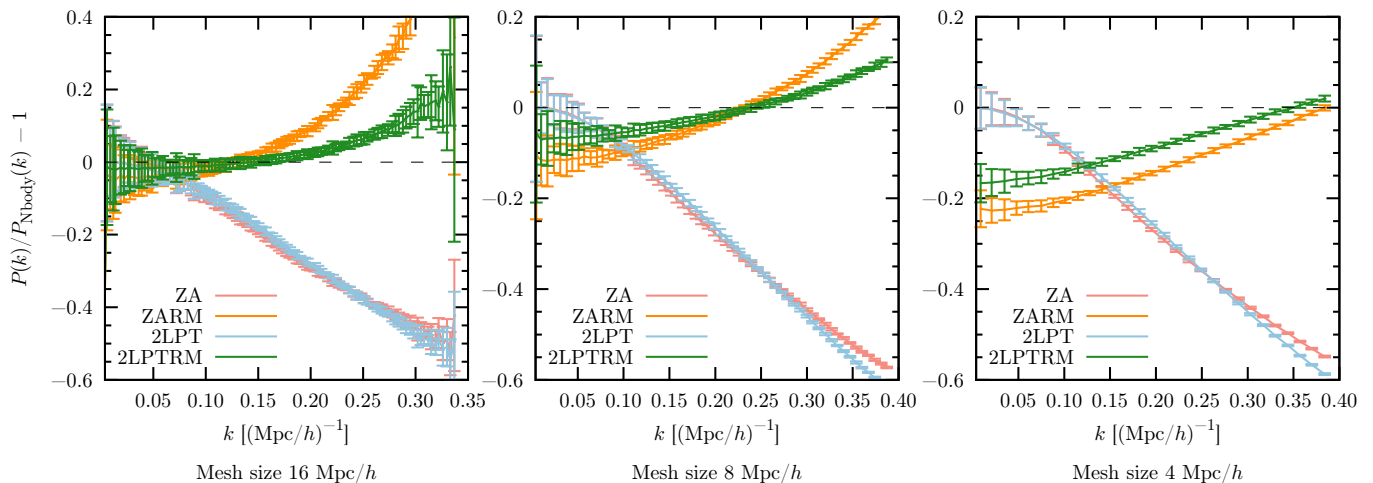
FIG. 9. *Power spectrum: mesh size-dependence.* Relative deviations for the power spectra of various particle distributions, with reference to the density field computed with a full N-body simulation. The particle distribution is determined using: the Zel'dovich approximation, alone (ZA, light red curve) and after remapping (ZARM, orange curve), second-order Lagrangian perturbation theory, alone (2LPT, light blue curve) and after remapping (2LPTRM, green curve). The computation is done on different meshes: 16 Mpc/h (64^3-voxel grid, left panel), 8 Mpc/h (128^3-voxel grid, central panel) and 4 Mpc/h (256^3-voxel grid, right panel). All results are shown at redshift $z = 0$. LPT fields exhibit more small-scale correlations after remapping and their power spectra get closer to the shape of the full non-linear power spectrum.
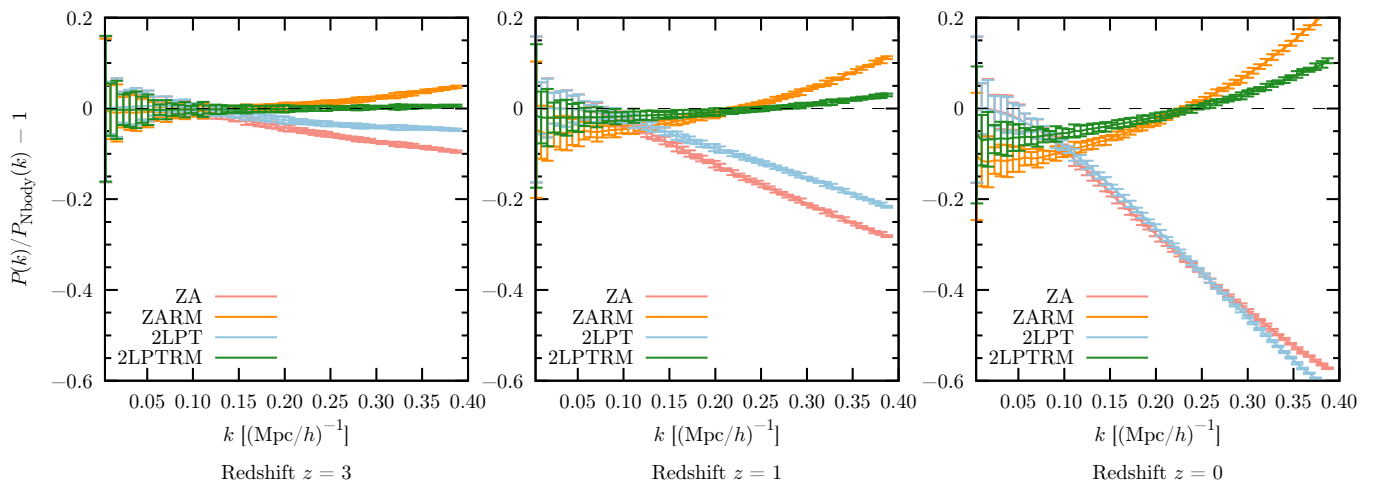

FIG. 10. *Power spectrum: redshift-dependence.* Relative deviations for the power spectra of various particle distributions (see the caption of figure 9), with reference to the density field computed with a full N-body simulation. The computation is done on a 8 Mpc/h mesh (128^3-voxel grid). Results at different redshifts are shown: $z = 3$ (right panel), $z = 1$ (central panel) and $z = 0$ (left panel). The remapping procedure is increasingly successful with increasing redshift.

scales, as can be seen in figure 8. This mesh size corresponds to the target resolution for analyses in the mildly non-linear regime, as discussed in the introduction.

2. Fourier-space cross-correlation coefficient

In figure 11, we present the Fourier-space cross-correlation coefficient $R(k) \equiv P_{\delta \times \delta'}/\sqrt{P_\delta P_{\delta'}}$ between the

redshift-zero density field in the N-body simulation and several other density fields. At this point, it is useful to recall that an approximation well-correlated with the non-linear density field can be used in a variety of cosmological applications, such as the reconstruction of the non-linear power spectrum (Tassev & Zaldarriaga, 2012b). As pointed out by Neyrinck (2013), the cross-correlation between 2LPT and full gravitational dynamics is higher at small k than the cross-correlation between the ZA and the full dynamics, meaning that the posi-
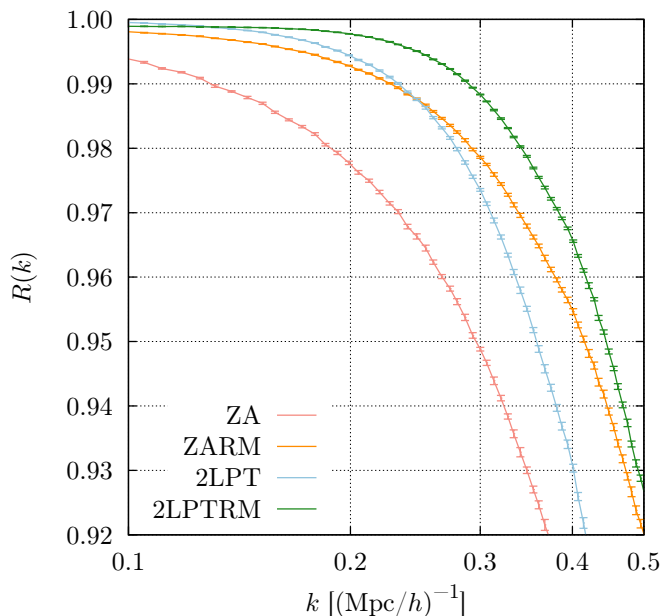
FIG. 11. Fourier-space cross-correlation coefficient between various approximately-evolved density fields and the particle distribution as evolved with full N-body dynamics, all at redshift zero. The binning of density fields is done on a 8 Mpc/h mesh (128^3-voxel grid). At small scales, $k \geq 0.2$ (Mpc/h)$^{-1}$, the cross-correlations with respect to the N-body-evolved field are notably better after remapping than with LPT alone.

tion of structures is more correct when additional physics (non-local tidal effects) are taken into account.

At redshift zero and at small scales, the agreement is better with remapped LPT fields than with LPT alone, confirming the success of the remapping procedure to explore the mildly non-linear regime. In particular, the remapping of 2LPT predicts more than 96% level accuracy at $k = 0.4$ (Mpc/h)$^{-1}$ (corresponding to scales of 16 Mpc/h), where 2LPT gives only 93%. The cross-correlation coefficient indicates better agreements for the remapping of 2LPT than for the remapping of the ZA, which is consistent with the better performance of 2LPT in predicting the phases of the full N-body field.

C. Three-point statistics

We analyzed the accuracy of our method beyond second-order statistics, by studying the three-point correlation function in Fourier space, the bispectrum. The importance of three-point statistics relies in their ability to test the correctness of the shape of structures. Some of the natural applications are to test gravity (Shirata *et al.*, 2007; Gil-Marín *et al.*, 2011), to break degeneracies in the galaxy bias (Matarrese, Verde & Heavens, 1997; Verde *et al.*, 1998; Scoccimarro *et al.*, 2001; Verde *et al.*, 2002) or to test the existence of primordial non-Gaussianities in the initial matter density field (Sefusatti
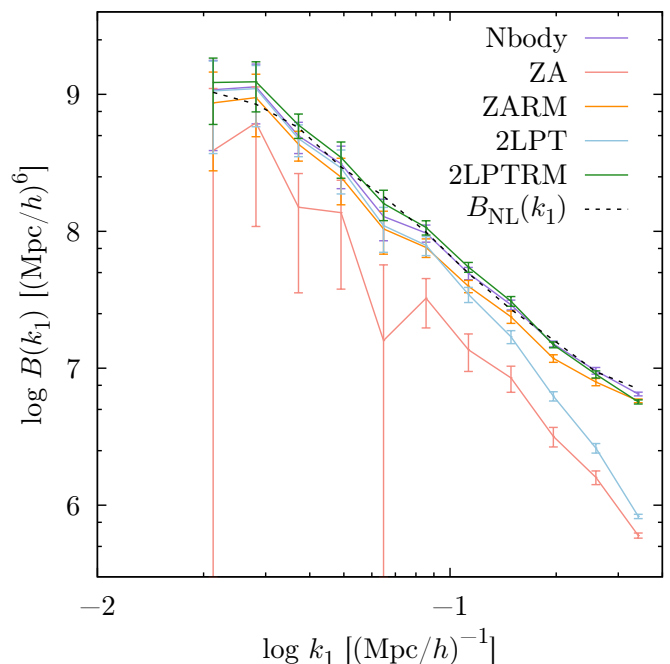

FIG. 12. Redshift-zero dark matter bispectra for equilateral triangle shape, in 1024 Mpc/h simulations, with density fields computed on mesh of 8 Mpc/h size. The particle distribution is determined using: a full N-body simulation (purple curve), the Zel'dovich approximation, alone (ZA, light red curve) and after remapping (ZARM, orange curve), second-order Lagrangian perturbation theory, alone (2LPT, light blue curve) and after remapping (2LPTRM, green curve). The dashed line, $B_{\mathrm{NL}}(k)$, corresponds to theoretical predictions for the bispectrum, found using the fitting formula of (Gil-Marín *et al.*, 2012). Note that both ZARM and 2LPTRM show increased bispectrum in the mildly non-linear regime compared to ZA and 2LPT, indicating an improvement of three-point statistics with the remapping procedure.

& Komatsu, 2007; Jeong & Komatsu, 2009).

As for the power spectrum, for the bispectrum we construct the dark matter density contrast field, namely $\delta(\mathbf{k})$, putting particles in cells using a CiC scheme, and deconvolving them after performing the Fourier transform.

The algorithm used to compute the bispectrum $B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ from this $\delta(\mathbf{k})$ field consists of randomly drawing k-vectors from a specified bin, namely Δk, and randomly orientating the $(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ triangle in space. We chose the number of random triangles to depend on the number of fundamental triangle per bin, that scales as $k_1 k_2 k_3 \Delta k^3$ (Scoccimarro, 1997), where Δk is the chosen k-binning: given k_i we allow triangles whose i-side lies between $k_i - \Delta k/2$ and $k_i + \Delta k/2$. In this paper we always set $\Delta k = k_{\min} = 2\pi/L$, where L is the size of the box. For the equilateral case, at scales of $k \approx 0.1$ (Mpc/h)$^{-1}$ we are generating $\sim 1.7 \times 10^6$ random triangles. We have verified that increasing the number of triangles beyond this value does not have any effect on the measurement.

As a rule of thumb, a maximum threshold in k for trusting the simulation data is set by a quarter of the Nyquist frequency, defined by $k_N = 2\pi/L \times N_v^{1/3}/2$, where L is the size of the box and N_v is the number of voxels in the grid on which particles are placed in the initial conditions; which makes for our analysis ($L = 1024$ Mpc/h, $N_v = 512^3$), $k_N/4 \approx 0.39$ (Mpc/h)$^{-1}$. At this scale, it has been observed that the power spectrum starts to deviate at the percent-level with respect to higher resolution simulations (Heitmann *et al.*, 2010). For all plots and results shown in this paper, this limit in k is never exceeded. Also, as a lower limit in k, we have observed that for scales larger than $\sim 3\,k_{min}$ the effects of sample variance start to play an important role on the bispectrum and considerable deviations with respect to linear theory are observed. Therefore the largest scale for the bispectrum analysis is set to $3\,k_{min} \approx 1.8 \times 10^{-2}$ (Mpc/h)$^{-1}$.

The error bars presented in the bispectrum plots represent the dispersion of the mean among eight independent realizations, all of them with the same cosmological parameters. It has been tested (Gil-Marín *et al.*, 2012), that this estimator for the error is in good agreement with theoretical predictions based on the Gaussianity of initial conditions (Scoccimarro, 1998).

Note that the subtracted shot noise is always assumed to be Poissonian:

$$B_{SN}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \frac{1}{\bar{n}}\left[P(k_1) + P(k_2) + P(k_3)\right] + \frac{1}{\bar{n}^2}, \tag{10}$$

(see e.g. Peebles, 1980, and references therein), where $\bar{n}$ is the number density of particles in the box.

Throughout this paper, a triangle shape is defined by the relative length of vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ and the inner angle θ_{12}, in such a way that $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0$ and $\mathbf{k}_1 \cdot \mathbf{k}_2 = k_1 k_2 \cos(\pi - \theta_{12})$. In figure 12, we plot the redshift-zero bispectrum, computed on a 8 Mpc/h mesh, of the different density fields for equilateral triangles ($\theta_{12} = \pi/3$ and $k_2/k_1 = 1$). There, the dashed line corresponds to theoretical predictions for the non-linear bispectrum, found using the fitting formula of Gil-Marín *et al.* (2012).

The overall result is a clear improvement of the bispectrum of LPT-evolved fields with the remapping procedure, especially on the small scales shown, probing the mildly non-linear regime, $k \gtrsim 0.1$ (Mpc/h)$^{-1}$ corresponding to scales $\lesssim 62$ Mpc/h, where LPT predicts less three-point correlation than full gravity. At large scales, the approximation error remains $\lesssim 1\sigma$ of the estimated statistical uncertainty, even for a resolution of 8 Mpc/h and at late times ($z = 0$).

The relative deviations of various bispectra with reference to full N-body simulations are shown in figures 13, 14, 15 and 16. As expected, the success of our remapping procedure in exploring small scales ($k \gtrsim 0.1$ (Mpc/h)$^{-1}$) is increased for more coarsely-binned density fields (see figure 13) and at higher redshift (see figure 14). In figure 15 we examine the scale-dependence of the bispectrum

for various triangle shapes. The precise dependence on the triangle shape at different scales is shown in figure 16.

IV. DISCUSSION AND CONCLUSION

The main subject of this paper is the development of a method designed to improve the correspondence between approximate models for gravitational dynamics and full numerical simulation of large-scale structure formation. Our methodology relies on a remapping of the one-point distribution of the density contrast of the approximately evolved particle distribution using information extracted from N-body simulations.

Due to the differences in the precise structure of the density fields in Lagrangian perturbation theory and in full gravity, the naive implementation of this procedure, inspired by W92, gives a large-scale bias in the power spectrum. This is not solved by a simple rescaling of Fourier modes, which leads to Gibbs ringing artifacts and an overall unphysical representation of the density fields. Smoothing LPT and N-body density fields with the same kernel is also unsuccessful, as smoothed N-body fields will always keep a sharper structure than smoothed LPT fields.

We figured out that the cause of this large-scale bias is not the different density predicted locally by LPT and N-body dynamics on a point-by-point basis, but a problem of mismatch between the volume of extended objects. Our findings question the reliability of LPT for LSS data analysis and generation of mock catalogs at low redshift and high mass resolution. They are also a likely explanation for the discrepancy between the power spectrum of initial conditions reconstructed via Gaussianization and linear theory expectations, encountered by W92.

Considering these results, we improved Weinberg's remapping procedure for present-day density fields by the use of a transfer function. In this fashion, we obtain a physically more accurate representation of the three-dimensional matter distribution in the mildly non-linear regime, while improving higher-order statistics. Since LPT captures well the cosmological dependence and remapping operates on small-scale aspects of the density field, we found that our procedure is nearly independent of cosmological parameters.

The aim of this method is to develop a fast, flexible and efficient way to generate realizations of LSS density fields, accurately representing the mildly non-linear regime. Our procedure, therefore, responds to the increasing demand for numerically inexpensive models of three-dimensional LSS, for applications to modern cosmological data analysis. At the level of statistical error in our numerical experiments, the approach provides a good method for producing mock halo catalogs and low-redshift initial conditions for simulations, if desired. The resulting information can also be used in a variety of cosmological analyses of present and upcoming observations.
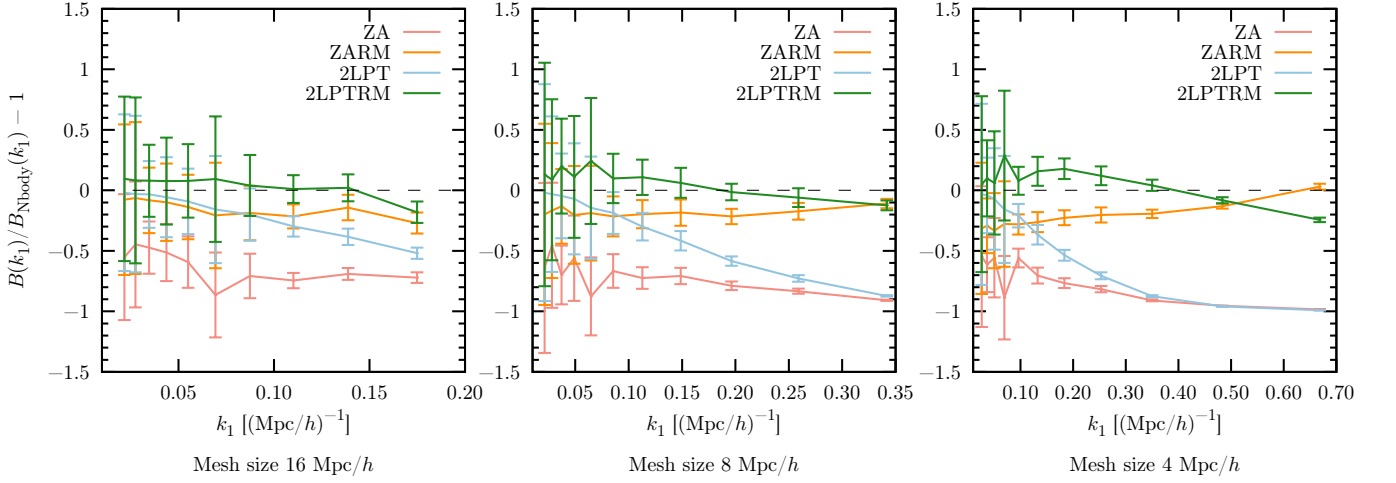
FIG. 13. *Bispectrum: mesh size-dependence.* Relative deviations for the bispectra $B(k_1)$ of various particle distributions, with reference to the prediction from a full N-body simulation, $B_{\mathrm{Nbody}}(k_1)$. The particle distribution is determined using: the Zel'dovich approximation, alone (ZA, light red curve) and after remapping (ZARM, orange curve), second-order Lagrangian perturbation theory, alone (2LPT, light blue curve) and after remapping (2LPTRM, green curve). The computation of bispectra is done for equilateral triangles and on different meshes: 16 Mpc/h (64^3-voxel grid, left panel), 8 Mpc/h (128^3-voxel grid, central panel) and 4 Mpc/h (256^3-voxel grid, right panel). All results are shown at redshift $z = 0$. LPT fields exhibit more small-scale three-point correlations after remapping and their bispectra get closer to the shape of the full non-linear bispectrum.
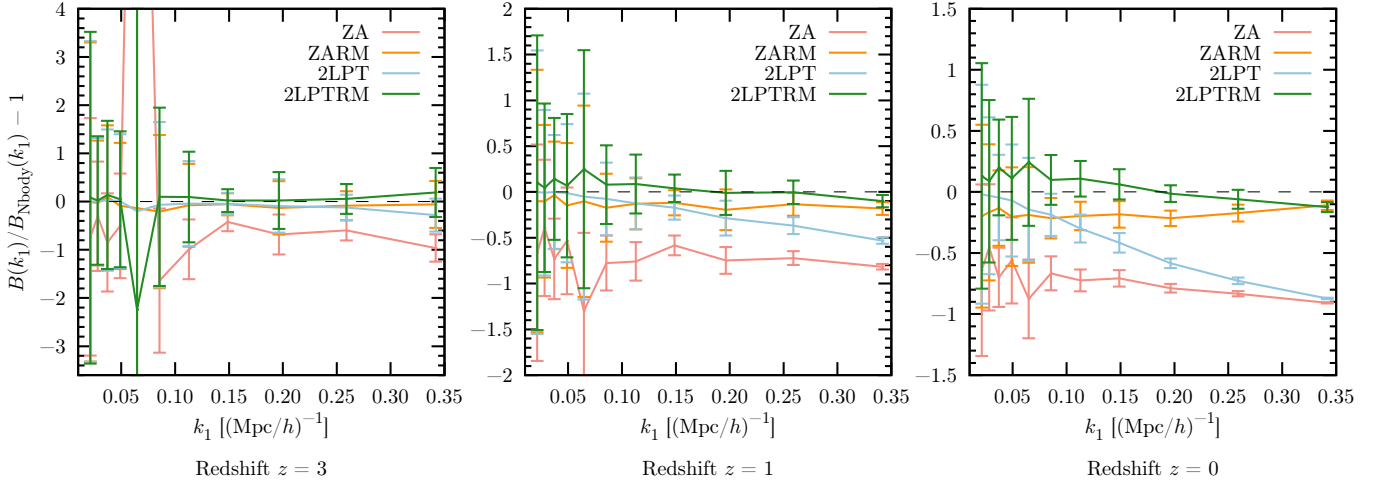

FIG. 14. *Bispectrum: redshift-dependence.* Relative deviations for the bispectra $B(k_1)$ of various particle distributions (see the caption of figure 13), with reference to a full N-body simulation, $B_{\mathrm{Nbody}}(k_1)$. The computation of bispectra is done on a 8 Mpc/h mesh (128^3-voxel grid) and for equilateral triangles. Results at different redshifts are shown: $z = 3$ (right panel), $z = 1$ (central panel) and $z = 0$ (left panel).

We showed that our approach allows fast generation of cosmological density fields that correlate with N-body simulations at better than 96% down to scales of $k \approx 0.4\,(\mathrm{Mpc}/h)^{-1}$ at redshift zero and are substantially better than standard LPT results at higher redshifts on the same comoving scales. Remapping improves the fast LPT bispectrum predictions on small scales while the large scale bispectrum remains accurate to within about 1σ of the measurement in our N-body simulations. Since

real observations will have larger statistical errors for the foreseeable future, our method provides an adequate fast model of the non-linear density field on scales down to $\sim$ 8 Mpc/h. These results constitute a substantial improvement with respect to existing techniques, since non-linearities begin to affect even large-scale measurements in galaxy surveys. Since the number of modes usable for cosmological exploitation scale as k^3, even minor improvements in the smallest scale k allow access to
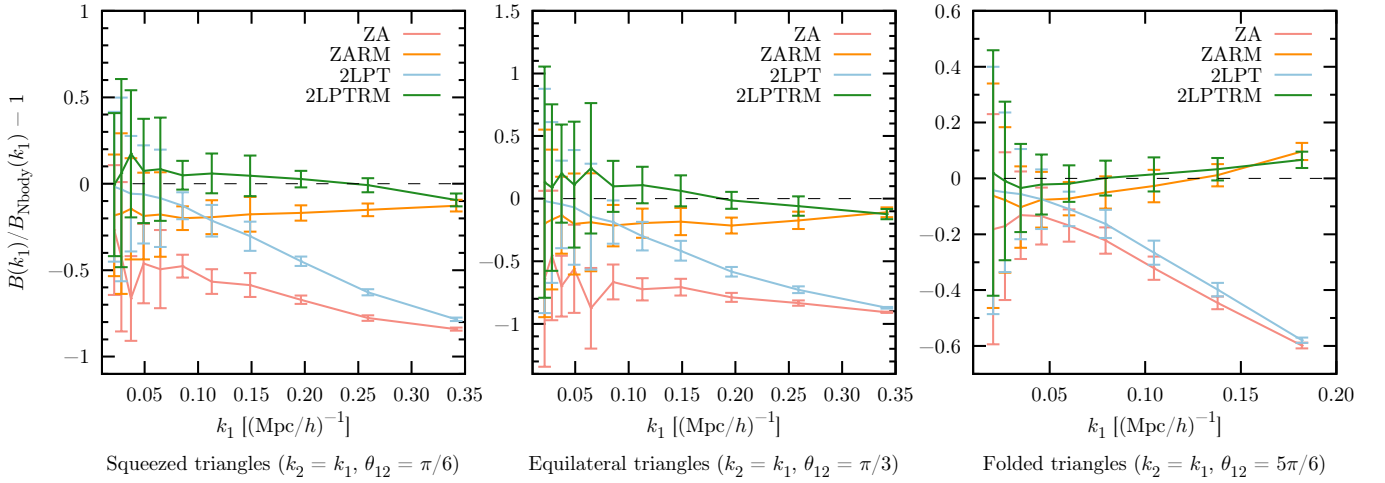
FIG. 15. *Bispectrum: scale-dependence for different triangle shapes.* Relative deviations for the bispectra $B(k_1)$ of various particle distributions (see the caption of figure 13), with reference to a full N-body simulation, $B_{\mathrm{Nbody}}(k_1)$. The computation is done on a 8 Mpc/h mesh (128^3-voxel grid) and results are shown at redshift $z = 0$ for various triangle shapes as defined above.
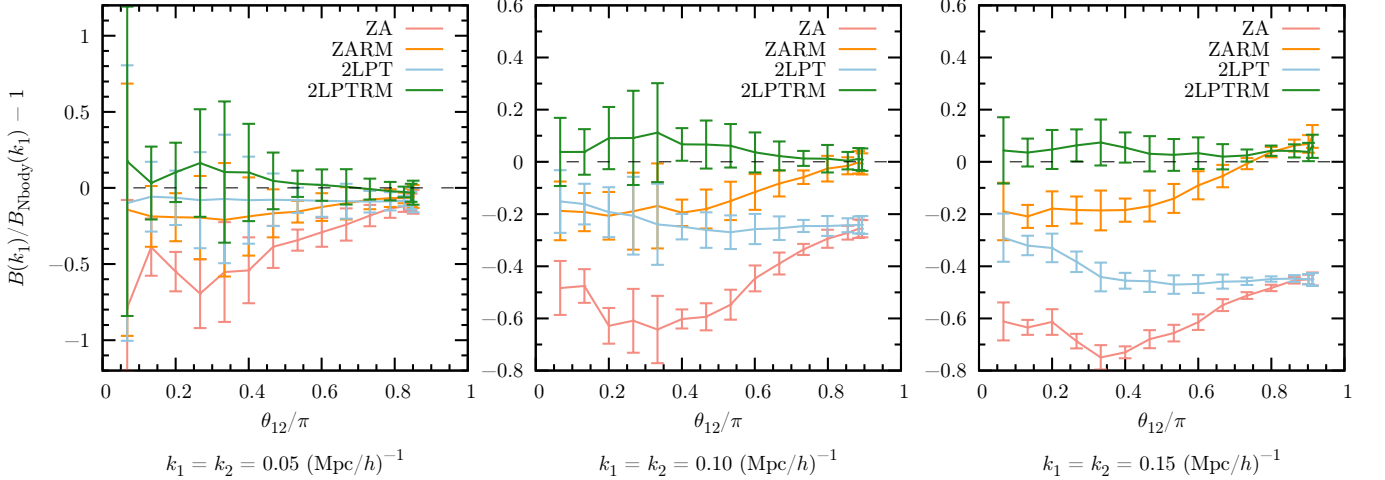

FIG. 16. *Bispectrum: triangle shape-dependence.* Relative deviations for the bispectra $B(k_1)$ of various particle distributions (see the caption of figure 13), with reference to a full N-body simulation, $B_{\mathrm{Nbody}}(k_1)$. The computation is done on a 8 Mpc/h mesh (128^3-voxel grid) and results are shown at redshift $z = 0$. The dependence on the angle of the triangle $\theta_{12} = (\mathbf{k}_1, \mathbf{k}_2)$ is shown for different scales: $k_1 = k_2 = 0.05$ (Mpc/h)$^{-1}$ (corresponding to 125 Mpc/h), $k_1 = k_2 = 0.10$ (Mpc/h)$^{-1}$ (corresponding to 63 Mpc/h), $k_1 = k_2 = 0.15$ (Mpc/h)$^{-1}$ (corresponding to 42 Mpc/h).

much more knowledge from existing and upcoming observations. This work is a step further in the non-linear regime, which contains a wealth of yet unexploited cosmological information. For possible applications, we provided a cosmographic and statistical characterization of approximation errors.

Our remapping procedure predicts the two-point correlation function at around 95% level accuracy and three-point correlation function at around 80% level accuracy at redshift 3, for k between 0.1 and 0.4 (Mpc/h)$^{-1}$, illustrating the increasing success of our methods as we go backwards in time towards the initial conditions, when

LPT is an accurate description of early structure formation. This is of particular interest in several areas of high-redshift cosmology, such as forecasting 21 cm surveys (Lidz *et al.*, 2007), analyzing the properties of the intergalactic medium via the Lyman-α forest (Kitaura, Gallerani & Ferrara, 2012) or probing the reionization epoch (Mesinger & Furlanetto, 2007). This work might also add to methods of data analysis of several ongoing

(Hетdex[1]) and upcoming high-redshift galaxy surveys (SDSS IV: eBOSS[2], DESI[3]).

However, the realization of density fields with these procedures stays approximate, since the full non-linear gravitational physics involves information contained in the shape of structures, which cannot be captured from a one-point modification of LPT, especially after shell-crossing. We studied the performance of one-point remapping of LPT and presented a statistical characterization of the errors, but additional refinements, such as a non-linear, density-dependent smoothing of the N-body field, could further improve on these approximations, for an increased computational cost. This is, however, beyond the scope and intent of this paper. Generally, the complications at large scales that we encounter when applying a local remapping seem difficult to solve in a Eulerian density field approach and would favor a Lagrangian, particle-based perspective, which will be a subject of further study.

Fast and accurate methods to model the non-linearly evolved mass distribution in the Universe have the potential of profound influence on modern cosmological data analysis. Recently proposed full Bayesian large scale structure inference methods, which extract information on the matter distribution in the Universe from galaxy redshift surveys, rely on Lagrangian perturbation theory (Jasche & Wandelt, 2013). The methods proposed in this paper will provide a numerically efficient and flexible extension of these methods, permitting us to push dynamic analyses of the large scale structure further into the non-linear regime. In particular, we envision our techniques to add to methods of physical inference and reconstruction of initial conditions from which the large-scale structure of the Universe originates. We may hope that this approach will yield insights into galaxy formation, a more detailed view on cosmic structure formation, on the nature of dark energy, as well as on physics of the early Universe by studying, for example, primordial non-Gaussianity via higher-order statistics.

Addendum: One-point statistics of the Lagrangian displacement field

The remapping procedure, described in section II, relies on the Eulerian density contrast. As noted by previous authors (see in particular Neyrinck, 2013), in the Lagrangian representation of the LSS, it is natural to use the divergence of the displacement field ψ instead of the Eulerian density contrast δ. This addendum provides additional comments on the one-point statistics of ψ and comparatively analyzes key features of ψ and δ.

———

[1] http://hetdex.org/
[2] http://www.sdss3.org/future/eboss.php
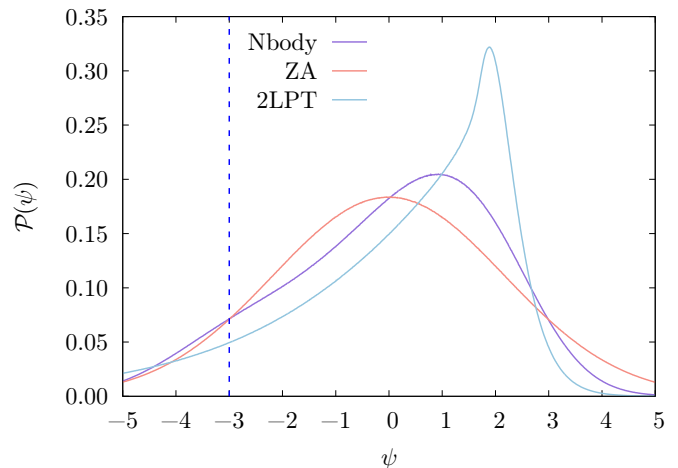[3] http://desi.lbl.gov/

FIG. 17. Redshift-zero probability distribution function for the divergence of the displacement field ψ, computed from eight 1024 Mpc/h-box simulations of 512^3 particles. A quantitative analysis of the deviation from Gaussianity of these PDFs is given in table I. The particle distribution is determined using: a full N-body simulation (purple curve), the Zel'dovich approximation (ZA, light red curve) and second-order Lagrangian perturbation theory (2LPT, light blue curve). The vertical line at $\psi = -3$ represents the collapse barrier about which ψ values bob around after gravitational collapse. A bump at this value is visible with full gravity, but LPT is unable to reproduce this feature. This regime corresponds to virialized, overdense clusters.

Model	$\mathcal{P}_\delta$	$\mathcal{P}_\psi$
	Skewness γ_1	
ZA	2.36 ± 0.01	-0.0067 ± 0.0001
2LPT	2.83 ± 0.01	-1.5750 ± 0.0002
N-body	5.14 ± 0.05	-0.4274 ± 0.0001
	Excess kurtosis γ_2	
ZA	9.95 ± 0.09	$-2.2154 \times 10^{-6} \pm 0.0003$
2LPT	13.91 ± 0.15	3.544 ± 0.0011
N-body	62.60 ± 2.75	-0.2778 ± 0.0004

TABLE I. Non-Gaussianity parameters (the skewness γ_1 and the excess kurtosis γ_2) of the redshift-zero probability distribution functions $\mathcal{P}_\delta$ and $\mathcal{P}_\psi$ of the density contrast δ and the divergence of the displacement field ψ, respectively. The confidence intervals given correspond to the 1-σ standard deviations among eight realizations. In all cases, γ_1 and γ_2 are reduced when measured from ψ instead of δ.

In the Lagrangian frame, the quantity of interest is not the position, but the displacement field $\Psi(\mathbf{q})$ which maps the initial comoving particle position $\mathbf{q}$ to its final comoving Eulerian position $\mathbf{x}$ (see e.g. Bouchet et al., 1995 or Bernardeau et al., 2002 for overviews),

$$\mathbf{x} \equiv \mathbf{q} + \Psi(\mathbf{q}). \qquad (A.11)$$

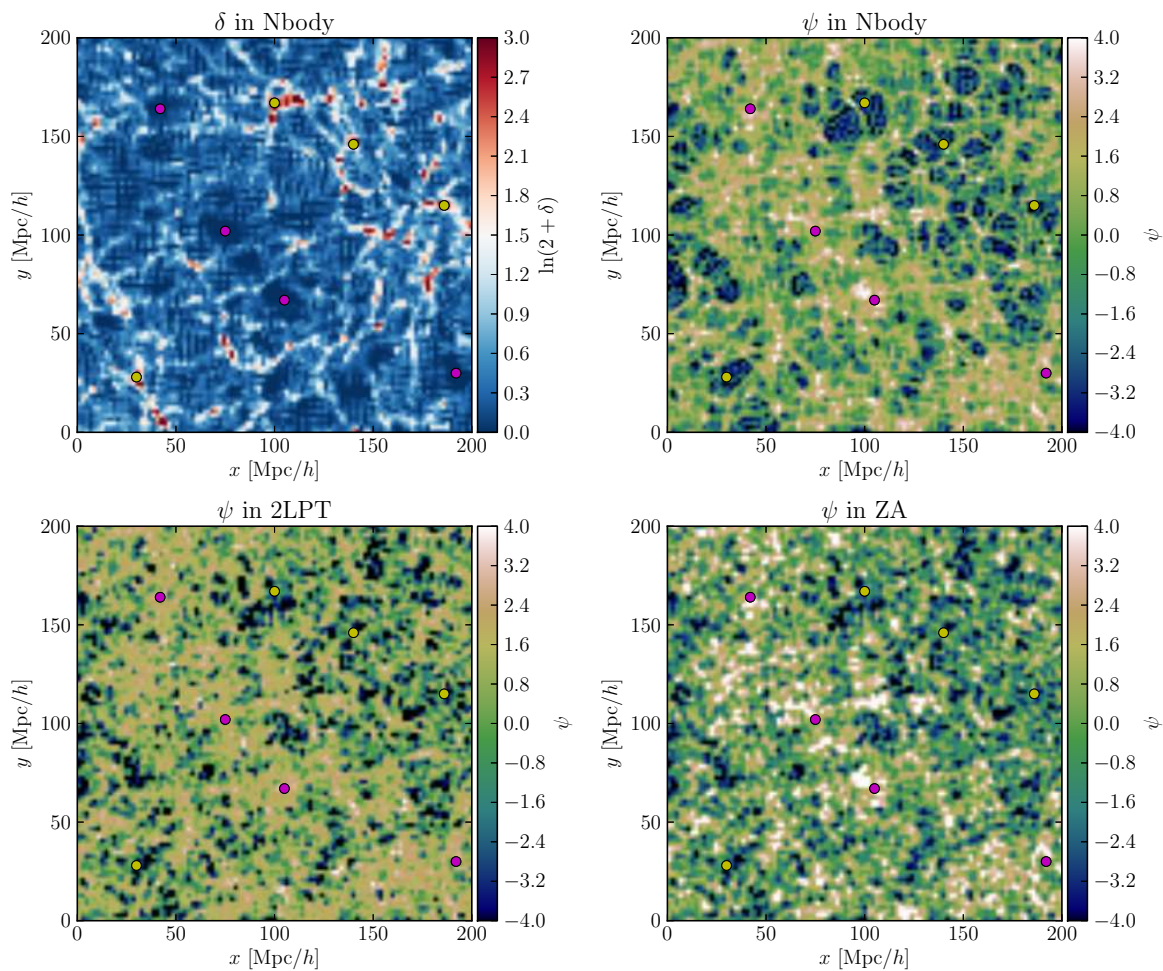It is important to note that, though $\Psi(\mathbf{q})$ is a priori a full

FIG. 18. Slices of the divergence of the displacement field, ψ, on a Lagrangian sheet of 512^2 particles from a 512^3-particle simulation of box size 1024 Mpc/h, run to redshift zero. For clarity we show only a 200 Mpc/h region. Each pixel corresponds to a particle. The particle distribution is determined using respectively a full N-body simulation, the Zel'dovich approximation (ZA) and second-order Lagrangian perturbation theory (2LPT). In the upper left panel, the density contrast δ in the N-body simulation is shown, after binning on a 512^3-voxel grid. To guide the eye, some clusters and voids are identified by yellow and purple dots, respectively. The "lakes", Lagrangian regions that have collapsed to form halos, are only visible in the N-body simulation, while the "mountains", Lagrangian regions corresponding to cosmic voids, are well reproduced by LPT.

three-dimensional vector field, it is curl-free up to second order in Lagrangian perturbation theory (appendix D in Bernardeau, 1994 or Bernardeau *et al.*, 2002 for a review). We did not consider contributions beyond 2LPT. After publication of the main paper, Chan (2014) analyzed the non-linear evolution of Ψ, splitting it into its scalar and vector parts (the so-called "Helmholtz decomposition"). Looking at two-point statistics, he found that shell-crossing leads to a suppression of small-scale power in the scalar part, and, subdominantly, to the generation of a vector contribution.

Let $\psi(\mathbf{q}) \equiv \nabla_\mathbf{q} \cdot \Psi(\mathbf{q})$ denote the divergence of the displacement field, where $\nabla_\mathbf{q}$ is the divergence operator in Lagrangian coordinates. ψ quantifies the angle-averaged spatial-stretching of the Lagrangian dark matter "sheet" in comoving coordinates (Neyrinck, 2013). Let $\mathcal{P}_{\psi,\mathrm{LPT}}$

and $\mathcal{P}_{\psi,\mathrm{Nbody}}$ be the one-point probability distribution functions for the divergence of the displacement field in LPT and in N-body fields, respectively. We denote by $\mathcal{P}_\delta$ the corresponding PDFs for the Eulerian density contrast.

In figure 17, we show the PDFs of ψ for the ZA, 2LPT and full N-body gravity. The most important feature of ψ is that, whatever the model for structure formation, the PDF exhibits reduced non-Gaussianity compared to the PDF for the density contrast δ (see the upper panel of figure 7 for comparison). The main reason is that $\mathcal{P}_\delta$, unlike $\mathcal{P}_\psi$, is tied down to zero at $\delta = -1$. It is highly non-Gaussian in the final conditions, both in N-body simulations and in approximations to the true dynamics. For a quantitative analysis, we looked at the first and second-order non-Gaussianity statistics: the skewness γ_1
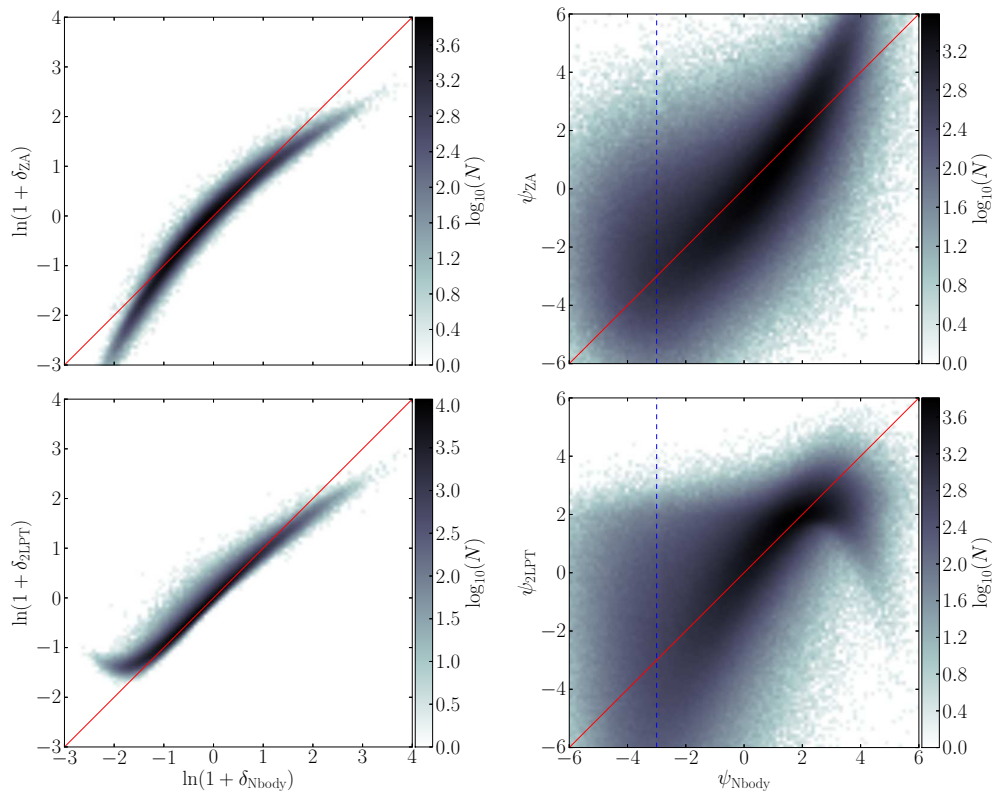
FIG. 19. *Left panel.* Two-dimensional histograms comparing particle densities evolved with full N-body dynamics (the x-axis) to densities in the LPT-evolved particle distributions (the y-axis). The red lines show the ideal $y = x$ locus. A turn-up at low densities is visible with 2LPT, meaning that some overdense regions are predicted where there should be deep voids. *Right panel.* Same plot for the divergence of the displacement field ψ. Negative ψ corresponds to overdensities and positive ψ correspond to underdensities. The dotted blue line shows the collapse barrier at $\psi = -3$ where particle get clustered in full gravity. The scatter is bigger with ψ than with δ, in particular in overdensities, since with LPT, particles do not cluster. The turn-up at low densities with 2LPT, observed with the density contrast, is also visible with the divergence of the displacement field.

and the excess kurtosis γ_2,

$$\gamma_1 \equiv \frac{\mu_3}{\sigma^3} \quad \text{and} \quad \gamma_2 \equiv \frac{\mu_4}{\sigma^4} - 3, \qquad \text{(A.12)}$$

where μ_n is the n-th moment about the mean and σ is the standard deviation. We estimated γ_1 and γ_2 at redshift zero in our simulations, in the one-point statistics of the density contrast δ and of the divergence of the displacement field ψ. The results are shown in table I. In all cases, we found that both γ_1 and γ_2 are much smaller when measured from $\mathcal{P}_\psi$ instead of $\mathcal{P}_\delta$.

At linear order in Lagrangian perturbation theory (the Zel'dovich approximation), the divergence of the displacement field is proportional to the density contrast in the initial conditions, $\delta(\mathbf{q})$, scaling with the negative growth factor, $-D_1(\tau)$:

$$\psi^{(1)}(\mathbf{q}, \tau) = \nabla_{\mathbf{q}} \cdot \Psi^{(1)}(\mathbf{q}, \tau) = -D_1(\tau)\delta(\mathbf{q}). \qquad \text{(A.13)}$$

Since we take Gaussian initial conditions, the PDF for ψ is Gaussian at any time with the ZA. In full gravity, non-linear evolution slightly breaks Gaussianity. $\mathcal{P}_{\psi,\text{Nbody}}$ is

slightly skewed towards negative values while its mode gets shifted around $\psi \approx 1$. Taking into account non-local effects, 2LPT tries to get closer to the shape observed in N-body simulations, but the correct skewness is overshot and the PDF is exceedingly peaked.

Figure 18 shows a slice of the divergence of the displacement field, measured at redshift zero for particles occupying a flat 512^2-pixel Lagrangian sheet from one of our simulations. For comparison, see also the figures in Mohayaee *et al.* (2006); Pueblas & Scoccimarro (2009); Neyrinck (2013). We used the color scheme of the latter paper, suggesting a topographical analogy when working in Lagrangian coordinates. As structures take shape, ψ departs from its initial value; it takes positive values in underdensities and negative values in overdensities. The shape of voids (the "mountains") is found to be reasonably similar in LPT and in the N-body simulation. For this reason, the influence of late-time non-linear effects in voids is milder as compared to overdense structures, which makes them easier to relate to the initial conditions. However, in overdense regions where ψ decreases,

it is not allowed to take arbitrary values: where gravitational collapse occurs, "lakes" form and ψ gets stuck around a collapse barrier, $\psi \approx -3$. As expected, these "lakes", corresponding to virialized clusters, can only be found in N-body simulations, since LPT fails to accurately describe the highly non-linear physics involved. A small bump at $\psi = -3$ is visible in $\mathcal{P}_{\psi,\text{Nbody}}$ (see figure 17). We checked that this bump gets more visible in higher mass-resolution simulations (200 Mpc/h box for 256^3 particles), where matter is more clustered. This means that part of the information about gravitational clustering can be found in the one-point statistics of ψ. Of course, the complete description of halos requires to precisely account for the shape of the "lakes", which can only be done via higher-order correlation functions. More generally, it is possible to use Lagrangian information in order to classify structures of the cosmic web. In particular, DIVA (Lavaux & Wandelt, 2010) uses the shear of the displacement field and ORIGAMI (Falck, Neyrinck & Szalay, 2012) the number of phase-space folds. As pointed out by Falck & Neyrinck (2015), while these techniques cannot be straightforwardly used for the analysis of galaxy surveys, where we lack Lagrangian information, recently proposed techniques for physical inference of the initial conditions (Jasche & Wandelt, 2013; Jasche, Leclercq & Wandelt, 2015) should allow their use with observational data.

Figure 19 shows two-dimensional histograms comparing N-body simulations to the LPT realizations for the density contrast δ and the divergence of the displacement field ψ. At this point, it is useful to note that a good mapping exists in the case where the relation shown is monotonic and the scatter is narrow. As pointed out by Sahni & Shandarin (1996) and Neyrinck (2013), matter in the substructure of 2LPT-voids has incorrect statistical properties: there are overdense particles in the low density region of the 2LPT δ-scatter plot. This degeneracy is also visible in the $\psi > 0$ region of the 2LPT ψ-scatter plot. On average, the scatter is bigger with ψ than with δ, in particular in overdensities ($\psi < 0$), since with LPT, particles do not cluster: ψ takes any value between 2 and -3 where it should remain around -3.

Summing up our discussions in this addendum, we analyzed the relative merits of the Lagrangian divergence of the displacement field ψ, and the Eulerian density contrast δ at the level of one-point statistics. The important differences are the following:

1. Ψ being irrotational up to order two, its divergence ψ contains nearly all information on the displacement field in one dimension, instead of three. The collapse barrier at $\psi = -3$ is visible in $\mathcal{P}_\psi$ for N-body simulations but not for LPT. A part of the information about non-linear gravitational clustering is therefore encoded in the one-point statistics of ψ.

2. ψ exhibits much fewer gravitationally-induced non-Gaussian features than δ in the final conditions (figure 17 and table I).

3. However, the values of ψ are more scattered than the values of δ with respect to the true dynamics (figure 19), meaning that an unambiguous mapping is more difficult.

Note added. While this addendum was being refeered, the work of Neyrinck (2015) appeared. It uses the spherical collapse prescription for ψ while checking various scales for the initial density field. The result is a fast scheme for producing approximate particle realizations.

(Bernardeau, 1994) F. Bernardeau, *The nonlinear evolution of rare events*, Astrophys. J. **427**, 51 (1994), astro-ph/9311066.

(Bernardeau *et al.*, 2002) F. Bernardeau, S. Colombi, E. Gaztañaga, R. Scoccimarro, *Large-scale structure of the Universe and cosmological perturbation theory*, Physics Reports **367**, 1 (2002), astro-ph/0112551.

(Bouchet *et al.*, 1995) F. R. Bouchet, S. Colombi, E. Hivon, R. Juszkiewicz, *Perturbative Lagrangian approach to gravitational instability*, Astron. & Astrophys. **296**, 575 (1995), astro-ph/9406013.

(Buchert, 1989) T. Buchert, *A class of solutions in Newtonian cosmology and the pancake theory*, Astron. & Astrophys. **223**, 9 (1989).

(Buchert, Melott & Weiss, 1994) T. Buchert, A. L. Melott, A. G. Weiss, *Testing higher-order Lagrangian perturbation theory against numerical simulations I. Pancake models*, Astron. & Astrophys. **288**, 349 (1994), astro-ph/9309056.

(Chan, 2014) K. C. Chan, *Helmholtz decomposition of the Lagrangian displacement*, Phys. Rev. D **89**, 083515 (2014), arXiv:1309.2243.

(Coles, Melott & Shandarin, 1993) P. Coles, A. L. Melott, S. F. Shandarin, *Testing approximations for non-linear gravitational clustering*, Mon. Not. R. Astron. Soc. **260**, 765 (1993).

(Crocce, Pueblas & Scoccimarro, 2006) M. Crocce, S. Pueblas, R. Scoccimarro, *Transients from initial conditions in cosmological simulations*, Mon. Not. R. Astron. Soc. **373**, 369 (2006), astro-ph/0606505.

(Crocce & Scoccimarro, 2006) M. Crocce, R. Scoccimarro, *Renormalized cosmological perturbation theory*, Phys. Rev. D **73**, 063519 (2006), astro-ph/0509418.

(Croft *et al.*, 1998) R. A. C. Croft, D. H. Weinberg, N. Katz, L. Hernquist, *Recovery of the Power Spectrum of Mass Fluctuations from Observations of the Ly alpha Forest*, Astrophys. J. **495**, 44 (1998), arXiv:astro-ph/9708018.

(Croft *et al.*, 1999) R. A. C. Croft, D. H. Weinberg, M. Pettini, L. Hernquist, N. Katz, *The Power Spectrum of Mass Fluctuations Measured from the LYalpha Forest at Redshift Z=2.5*, Astrophys. J. **520**, 1 (1999), arXiv:astro-ph/9809401.

(Doroshkevich, 1970) A. G. Doroshkevich, *The space structure of perturbations and the origin of rotation of galaxies in the theory of fluctuation.*, Astrofizika **6**, 581 (1970).

(Eisenstein *et al.*, 2005) D. J. Eisenstein, I. Zehavi, D. W. Hogg, R. Scoccimarro, M. R. Blanton, R. C. Nichol, R. Scranton, H.-J. Seo, M. Tegmark, Z. Zheng, S. F. Anderson, J. Annis, N. Bahcall, J. Brinkmann, S. Burles, F. J. Castander, A. Connolly, I. Csabai, M. Doi, M. Fukugita, J. A. Frieman, K. Glazebrook, J. E. Gunn, J. S. Hendry, G. Hennessy, Z. Ivezić, S. Kent, G. R. Knapp, H. Lin, Y.-S. Loh, R. H. Lupton, B. Margon, T. A. McKay, A. Meiksin, J. A. Munn, A. Pope, M. W. Richmond, D. Schlegel, D. P. Schneider, K. Shimasaku, C. Stoughton, M. A. Strauss, M. SubbaRao, A. S. Szalay, I. Szapudi, D. L. Tucker, B. Yanny, D. G. York, *Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies*, Astrophys. J. **633**, 560 (2005), astro-ph/0501171.

(Falck & Neyrinck, 2015) B. Falck, M. C. Neyrinck, *The persistent percolation of single-stream voids*, Mon. Not. R. Astron. Soc. **450**, 3239 (2015), arXiv:1410.4751.

(Falck, Neyrinck & Szalay, 2012) B. L. Falck, M. C. Neyrinck, A. S. Szalay, *ORIGAMI: Delineating Halos Using Phase-space Folds*, Astrophys. J. **754**, 126 (2012), arXiv:1201.2353 [astro-ph.CO].

(Feng & Fang, 2000) L.-L. Feng, L.-Z. Fang, *Non-Gaussianity and the Recovery of the Mass Power Spectrum from the Lyα Forest*, Astrophys. J. **535**, 519 (2000), arXiv:astro-ph/0001348.

(Gil-Marín *et al.*, 2011) H. Gil-Marín, F. Schmidt, W. Hu, R. Jimenez, L. Verde, *The bispectrum of f(R) cosmologies*, Journal of Cosmology and Astroparticle Physics **11**, 019 (2011), arXiv:1109.2115 [astro-ph.CO].

(Gil-Marín *et al.*, 2012) H. Gil-Marín, C. Wagner, F. Fragkoudi, R. Jimenez, L. Verde, *An improved fitting formula for the dark matter bispectrum*, Journal of Cosmology and Astroparticle Physics **2**, 047 (2012), arXiv:1111.4477 [astro-ph.CO].

(Gunn & Gott, 1972) J. E. Gunn, J. R. Gott, III, *On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution*, Astrophys. J. **176**, 1 (1972).

(Gurbatov, Saichev & Shandarin, 1989) S. N. Gurbatov, A. I. Saichev, S. F. Shandarin, *The large-scale structure of the universe in the frame of the model equation of non-linear diffusion*, Mon. Not. R. Astron. Soc. **236**, 385 (1989).

(Hahn *et al.*, 2007) O. Hahn, C. Porciani, C. M. Carollo, A. Dekel, *Properties of dark matter haloes in clusters, filaments, sheets and voids*, Mon. Not. R. Astron. Soc. **375**, 489 (2007), astro-ph/0610280.

(Heisenberg, Schäfer & Bartelmann, 2011) L. Heisenberg, B. M. Schäfer, M. Bartelmann, *A study of relative velocity statistics in Lagrangian perturbation theory with PINOCCHIO*, Mon. Not. R. Astron. Soc. **416**, 3057 (2011), arXiv:1011.1559 [astro-ph.CO].

(Heitmann *et al.*, 2009) K. Heitmann, D. Higdon, M. White, S. Habib, B. J. Williams, E. Lawrence, C. Wagner, *The Coyote Universe. II. Cosmological Models and Precision Emulation of the Nonlinear Matter Power Spectrum*, Astrophys. J. **705**, 156 (2009), arXiv:0902.0429 [astro-ph.CO].

(Heitmann *et al.*, 2010) K. Heitmann, M. White, C. Wagner, S. Habib, D. Higdon, *The Coyote Universe. I. Precision Determination of the Nonlinear Matter Power Spectrum*, Astrophys. J. **715**, 104 (2010), arXiv:0812.1052.

(Jasche, Leclercq & Wandelt, 2015) J. Jasche, F. Leclercq, B. D. Wandelt, *Past and present cosmic structure in the SDSS DR7 main sample*, Journal of Cosmology and Astroparticle Physics **1**, 036 (2015), arXiv:1409.6308.

(Jasche & Wandelt, 2013) J. Jasche, B. D. Wandelt, *Bayesian physical reconstruction of initial conditions from large-scale structure surveys*, Mon. Not. R. Astron. Soc. **432**, 894 (2013), arXiv:1203.3639 [astro-ph.CO].

(Jeong & Komatsu, 2009) D. Jeong, E. Komatsu, *Primordial Non-Gaussianity, Scale-dependent Bias, and the Bispectrum of Galaxies*, Astrophys. J. **703**, 1230 (2009), arXiv:0904.0497 [astro-ph.CO].

(Jing, 2005) Y. P. Jing, *Correcting for the Alias Effect When Measuring the Power Spectrum Using a Fast Fourier Transform*, Astrophys. J. **620**, 559 (2005), astro-ph/0409240.

(Kitaura, Gallerani & Ferrara, 2012) F.-S. Kitaura, S. Gallerani, A. Ferrara, *Multiscale inference of matter fields and baryon acoustic oscillations from the Lyα forest*, Mon. Not. R. Astron. Soc. **420**, 61 (2012), arXiv:1011.6233 [astro-ph.CO].

(Komatsu *et al.*, 2011) E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. R. Nolta, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, E. L. Wright, *Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation*, Astrophys. J. Supp. **192**, 18 (2011), arXiv:1001.4538 [astro-ph.CO].

(Lavaux & Wandelt, 2010) G. Lavaux, B. D. Wandelt, *Precision cosmology with voids: definition, methods, dynamics*, Mon. Not. R. Astron. Soc. **403**, 1392 (2010), arXiv:0906.4101 [astro-ph.CO].

(Lawrence *et al.*, 2010) E. Lawrence, K. Heitmann, M. White, D. Higdon, C. Wagner, S. Habib, B. Williams, *The Coyote Universe. III. Simulation Suite and Precision Emulator for the Nonlinear Matter Power Spectrum*, Astrophys. J. **713**, 1322 (2010), arXiv:0912.4490 [astro-ph.CO].

(Lidz *et al.*, 2007) A. Lidz, O. Zahn, M. McQuinn, M. Zaldarriaga, S. Dutta, L. Hernquist, *Higher Order Contributions to the 21 cm Power Spectrum*, Astrophys. J. **659**, 865 (2007), astro-ph/0610054.

(Manera *et al.*, 2013) M. Manera, R. Scoccimarro, W. J. Percival, L. Samushia, C. K. McBride, A. J. Ross, R. K. Sheth, M. White, B. A. Reid, A. G. Sánchez, R. de Putter, X. Xu, A. A. Berlind, J. Brinkmann, C. Maraston, B. Nichol, F. Montesano, N. Padmanabhan, R. A. Skibba, R. Tojeiro, B. A. Weaver, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: a large sample of mock galaxy catalogues*, Mon. Not. R. Astron. Soc. **428**, 1036 (2013), arXiv:1203.6609 [astro-ph.CO].

(Matarrese, Verde & Heavens, 1997) S. Matarrese, L. Verde, A. F. Heavens, *Large-scale bias in the Universe: bispectrum method*, Mon. Not. R. Astron. Soc. **290**, 651 (1997), astro-ph/9706059.

(Melott, 1993) A. L. Melott, *Improving the reconstruction of the velocity potential and primordial density fluctuations by choice of smoothing windows*, Astrophys. J. **414**, L73 (1993).

(Melott, Pellman & Shandarin, 1994) A. L. Melott, T. F. Pellman, S. F. Shandarin, *Optimizing the Zeldovich Approximation*, Mon. Not. R. Astron. Soc. **269**, 626 (1994), astro-ph/9312044.

(Mesinger & Furlanetto, 2007) A. Mesinger, S. Furlanetto, *Efficient Simulations of Early Structure Formation and Reionization*, Astrophys. J. **669**, 663 (2007), arXiv:0704.0946.

(Mohayaee *et al.*, 2006) R. Mohayaee, H. Mathis, S. Colombi, J. Silk, *Reconstruction of primordial density fields*, Mon. Not. R. Astron. Soc. **365**, 939 (2006), astro-ph/0501217.

(Monaco *et al.*, 2013) P. Monaco, E. Sefusatti, S. Borgani,

M. Crocce, P. Fosalba, R. K. Sheth, T. Theuns, *An accurate tool for the fast generation of dark matter halo catalogues*, Mon. Not. R. Astron. Soc. **433**, 2389 (2013), arXiv:1305.1505 [astro-ph.CO].

(Monaco, Theuns & Taffoni, 2002) P. Monaco, T. Theuns, G. Taffoni, *The pinocchio algorithm: pinpointing orbit-crossing collapsed hierarchical objects in a linear density field*, Mon. Not. R. Astron. Soc. **331**, 587 (2002), astro-ph/0109323.

(Monaco *et al.*, 2002) P. Monaco, T. Theuns, G. Taffoni, F. Governato, T. Quinn, J. Stadel, *Predicting the Number, Spatial Distribution, and Merging History of Dark Matter Halos*, Astrophys. J. **564**, 8 (2002), astro-ph/0109322.

(Moutarde *et al.*, 1991) F. Moutarde, J.-M. Alimi, F. R. Bouchet, R. Pellat, A. Ramani, *Precollapse scale invariance in gravitational instability*, Astrophys. J. **382**, 377 (1991).

(Munshi & Starobinsky, 1994) D. Munshi, A. A. Starobinsky, *Nonlinear approximations to gravitational instability: A comparison in second-order perturbation theory*, Astrophys. J. **428**, 433 (1994), astro-ph/9311056.

(Narayanan & Weinberg, 1998) V. K. Narayanan, D. H. Weinberg, *Reconstruction Analysis of Galaxy Redshift Surveys: A Hybrid Reconstruction Method*, Astrophys. J. **508**, 440 (1998), arXiv:astro-ph/9806238.

(Neyrinck, 2013) M. C. Neyrinck, *Quantifying distortions of the Lagrangian dark-matter mesh in cosmology*, Mon. Not. R. Astron. Soc. **428**, 141 (2013), arXiv:1204.1326 [astro-ph.CO].

(Neyrinck, 2015) M. C. Neyrinck, *Truthing the stretch: Nonperturbative cosmological realizations with multiscale spherical collapse*, ArXiv e-prints (2015), arXiv:1503.07534.

(Neyrinck, Szapudi & Szalay, 2011) M. C. Neyrinck, I. Szapudi, A. S. Szalay, *Rejuvenating Power Spectra. II. The Gaussianized Galaxy Density Field*, Astrophys. J. **731**, 116 (2011), arXiv:1009.5680 [astro-ph.CO].

(Neyrinck & Yang, 2013) M. C. Neyrinck, L. F. Yang, *Ringing the initial Universe: the response of overdensity and transformed-density power spectra to initial spikes*, Mon. Not. R. Astron. Soc. **433**, 1628 (2013), arXiv:1305.1629 [astro-ph.CO].

(Peebles, 1980) P. J. E. Peebles, *Research supported by the National Science Foundation. Princeton, N.J., Princeton University Press, 1980. 435 p.* (1980).

(Pueblas & Scoccimarro, 2009) S. Pueblas, R. Scoccimarro, *Generation of vorticity and velocity dispersion by orbit crossing*, Phys. Rev. D **80**, 043504 (2009), arXiv:0809.4606.

(Sahni & Shandarin, 1996) V. Sahni, S. Shandarin, *Accuracy of Lagrangian approximations in voids*, Mon. Not. R. Astron. Soc. **282**, 641 (1996), astro-ph/9510142.

(Scoccimarro, 1997) R. Scoccimarro, *Cosmological Perturbations: Entering the Nonlinear Regime*, Astrophys. J. **487**, 1 (1997), astro-ph/9612207.

(Scoccimarro, 1998) R. Scoccimarro, *Transients from initial conditions: a perturbative analysis*, Mon. Not. R. Astron. Soc. **299**, 1097 (1998), astro-ph/9711187.

(Scoccimarro, 2000) R. Scoccimarro, *The Bispectrum: From Theory to Observations*, Astrophys. J. **544**, 597 (2000), astro-ph/0004086.

(Scoccimarro *et al.*, 2001) R. Scoccimarro, H. A. Feldman, J. N. Fry, J. A. Frieman, *The Bispectrum of IRAS Redshift Catalogs*, Astrophys. J. **546**, 652 (2001), astro-ph/0004087.

(Scoccimarro & Sheth, 2002) R. Scoccimarro, R. K. Sheth, *PTHALOS: a fast method for generating mock galaxy distributions*, Mon. Not. R. Astron. Soc. **329**, 629 (2002), astro-ph/0106120.

(Sefusatti & Komatsu, 2007) E. Sefusatti, E. Komatsu, *Bispectrum of galaxies from high-redshift galaxy surveys: Primordial non-Gaussianity and nonlinear galaxy bias*, Phys. Rev. D **76**, 083004 (2007), arXiv:0705.0343.

(Shandarin & Zel'dovich, 1989) S. F. Shandarin, Y. B. Zel'dovich, *The large-scale structure of the universe: Turbulence, intermittency, structures in a self-gravitating medium*, Reviews of Modern Physics **61**, 185 (1989).

(Shirata *et al.*, 2007) A. Shirata, Y. Suto, C. Hikage, T. Shiromizu, N. Yoshida, *Galaxy clustering constraints on deviations from Newtonian gravity at cosmological scales. II. Perturbative and numerical analyses of power spectrum and bispectrum*, Phys. Rev. D **76**, 044026 (2007), arXiv:0705.1311.

(Springel, 2005) V. Springel, *The cosmological simulation code GADGET-2*, Mon. Not. R. Astron. Soc. **364**, 1105 (2005), astro-ph/0505010.

(Springel, Yoshida & White, 2001) V. Springel, N. Yoshida, S. D. M. White, *GADGET: a code for collisionless and gasdynamical cosmological simulations*, New Astronomy **6**, 79 (2001), astro-ph/0003162.

(Taffoni, Monaco & Theuns, 2002) G. Taffoni, P. Monaco, T. Theuns, *PINOCCHIO and the hierarchical build-up of dark matter haloes*, Mon. Not. R. Astron. Soc. **333**, 623 (2002), astro-ph/0109324.

(Tassev & Zaldarriaga, 2012a) S. Tassev, M. Zaldarriaga, *Estimating CDM particle trajectories in the mildly non-linear regime of structure formation. Implications for the density field in real and redshift space*, Journal of Cosmology and Astroparticle Physics **12**, 011 (2012a), arXiv:1203.5785 [astro-ph.CO].

(Tassev & Zaldarriaga, 2012b) S. Tassev, M. Zaldarriaga, *The mildly non-linear regime of structure formation*, Journal of Cosmology and Astroparticle Physics **4**, 013 (2012b), arXiv:1109.4939 [astro-ph.CO].

(Tassev & Zaldarriaga, 2012c) S. Tassev, M. Zaldarriaga, *Towards an optimal reconstruction of baryon oscillations*, Journal of Cosmology and Astroparticle Physics **10**, 006 (2012c), arXiv:1203.6066 [astro-ph.CO].

(Tassev, Zaldarriaga & Eisenstein, 2013) S. Tassev, M. Zaldarriaga, D. J. Eisenstein, *Solving large scale structure in ten easy steps with COLA*, Journal of Cosmology and Astroparticle Physics **6**, 036 (2013), arXiv:1301.0322 [astro-ph.CO].

(Valageas, 2007) P. Valageas, *Large-N expansions applied to gravitational clustering*, Astron. & Astrophys. **465**, 725 (2007), astro-ph/0611849.

(Verde *et al.*, 1998) L. Verde, A. F. Heavens, S. Matarrese, L. Moscardini, *Large-scale bias in the Universe - II. Redshift-space bispectrum*, Mon. Not. R. Astron. Soc. **300**, 747 (1998), astro-ph/9806028.

(Verde *et al.*, 2002) L. Verde, A. F. Heavens, W. J. Percival, S. Matarrese, C. M. Baugh, J. Bland-Hawthorn, T. Bridges, R. Cannon, S. Cole, M. Colless, C. Collins, W. Couch, G. Dalton, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, C. S. Frenk, K. Glazebrook, C. Jackson, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, D. Madgwick, P. Norberg, J. A. Peacock, B. A. Peterson, W. Sutherland, K. Taylor, *The 2dF Galaxy Redshift Survey: the bias of galaxies and the density of the Universe*, Mon. Not. R. Astron. Soc. **335**, 432 (2002), arXiv:astro-ph/0112161.

(Weinberg, 1992) D. H. Weinberg, *Reconstructing primordial density fluctuations. I - Method*, Mon. Not. R. Astron. Soc. **254**, 315 (1992).

(Yoshisato *et al.*, 2006) A. Yoshisato, M. Morikawa, N. Gouda, H. Mouri, *Why is the Zel'dovich Approximation So Accurate?*, Astrophys. J. **637**, 555 (2006), astro-ph/0510107.

(Yu *et al.*, 2011) Y. Yu, P. Zhang, W. Lin, W. Cui, J. N. Fry, *Gaussianizing the non-Gaussian lensing convergence field: The performance of the Gaussianization*, Phys. Rev. D **84**, 023523 (2011), arXiv:1103.2858 [astro-ph.CO].

(Yu *et al.*, 2012) Y. Yu, P. Zhang, W. Lin, W. Cui, J. N. Fry, *Gaussianizing the non-Gaussian lensing convergence field II. The applicability to noisy data*, Phys. Rev. D **86**, 023515 (2012), arXiv:1201.4527 [astro-ph.CO].

(Zel'dovich, 1970) Y. B. Zel'dovich, *Gravitational instability: An approximate theory for large density perturbations.*, Astron. & Astrophys. **5**, 84 (1970).